

COPY 82 OF 100

7.9-1007.0
TM-79933
LACIE-C00200

VOLUME II (REV. E)
JSC-11330

LARGE AREA CROP INVENTORY EXPERIMENT (LACIE)



"Made available under NASA sponsorship
in the interest of early and wide dis-
semination of Earth Resources Survey
Program information and without liability
for any use made thereof."

NASA NOAA USDA

N79-15349

(E79-10070) LARGE AREA CROP INVENTORY
EXPERIMENT (LACIE). LEVEL 3 BASELINE;
CLASSIFICATION AND MENSURATION SUBSYSTEM
(CAMS) REQUIREMENTS, VOLUME 2, REVISION E
(NASA) 121 P HC A06/MF A01 CSCL 02C G3/43

Unclass
00070

CLASSIFICATION AND MENSURATION SUBSYSTEM (CAMS) REQUIREMENTS

NASA

National Aeronautics and
Space Administration

Lyndon B Johnson Space Center
Houston, Texas 77058

MAY 1978

NOTICE THIS IS A BASELINED LEVEL 3 DOCUMENT CON-
TROLLED BY THE LACIE LEVEL 3 CHANGE CONTROL BOARD
ANY PROPOSED CHANGES SHOULD BE DOCUMENTED ON AN
RECP FORM AND TRANSMITTED TO R B MACDONALD,
LACIE MANAGER, NASA-JSC, CODE SF, HOUSTON, TEXAS 77058

REVISIONS

REV LTR	CHANGE NO.	DESCRIPTION	DATE
		BASELINE ISSUE (Reference CCBD #III-0001, dated December 16, 1974)*	12-16-74
A	5M0051	Major revision due to project rescoping	11-04-75
B	6M0029	Major revision due to extensive changes (LCN's 6M0006, 6M0007, 6M0009, 6M0013, 6M0015, 6M0022, 6M0005A are incorporated in this revision.)	7-14-76
C	6M0046	Major revision due to extensive changes (LCN's 6M0008, 6M0033, 6M0034A, 6M0036A, 6M0047, 6M0048, 6M0049, 6M0050, 6M0051 are incorporated in this revision.)	12-22-76
D	7M0012	Major revision due to extensive changes (LCN's 7M0013 through 7M0024, and 7M0026 are incorporated in this revision.)	7-21-77
E	8M0004	Major revision due to extensive changes (LCN's 8M0005 through 8M0009 are incorporated in this revision.)	5-31-78
*The changes required by the following RID's which were approved during the LACIE Project Review conducted December 3 through 5, 1974, have been incorporated into this baseline issue of the LACIE Level III Requirements Documents: 0-2 through 0-13, 0-15 through 0-29, and 0-31 through 0-46 1a-1, 1a-2, 1a-10, 1a-21, 1a-34, and 1a-36 1b-3 and 1b-6 1c-2, 1c-4, 1c-5, 1c-9, 1c-10, 1c-13, 1c-16, 1c-17, and 1c-20 2-12a, 2-13, 2-17, and 2-27 All other changes required by the remaining RID's approved during the Project Review will be incorporated by transmitting an RECP to the LACIE Level III Change Control Board for approval. Each RECP should be accompanied by the appropriate RID Closeout Form as described during the Project Review.			

LIST OF EFFECTIVE PAGES

The current status of all pages in this document is as shown below:

<u>Page No.</u>	<u>LACIE Change Date</u>	<u>Authorizing CCBD No.</u>
<u>ii</u> through <u>xi</u>	Revision E 5/31/78	8M0004
<u>1-1</u> through <u>A-23</u>	Revision E 5/31/78	8M0004

FOREWORD

Efficient management of the Large Area Crop Inventory Experiment (LACIE) dictates that effective controls of project activities be established. To provide a basis for effective control, documentation will be prepared, baselines will be established, and changes to the baseline will be subsequently controlled by the proper management levels.

The specific control documents which will be used are defined in the LACIE Project Plan, Document Number LACIE-C00605. All elements of the LACIE project must adhere to these baselined control documents, and, where it is considered that the requirements should be changed, the proper change request accompanied by a full justification must be submitted to the proper management level in accordance with established procedures. These documents will be maintained current by change notices and revisions as required. Each change notice and/or revision will reference the applicable Change Control Board Directive (CCBD) which approved the change.

This document, LACIE-C00200, Volume II, Revision E, defines the LACIE Classification and Mensuration Subsystem (CAMS) requirements and has been prepared in accordance with the "Instructions for Preparation of LACIE Requirements Documents", LACIE-00100, Revision C, dated November 20, 1974. "Full-Up System" as used in this document is defined as the system required to accomplish LACIE Phase II. In general, the approach used in each section is to first specify the requirements of the "Full-Up System" and then to specify the requirements of any interim systems by reference to specific paragraphs in the Full-Up System requirements sections of the document. The LACIE Project Phases are defined in the LACIE Project Plan, LACIE-C00605.

The organization responsible for the implementation of each requirement defined in this document is specified on an individual requirement basis. Where the implementation responsibility applies to the complete section, the implementation responsibility is specified after the section title. A "section" for the purposes of designating implementation responsibility is defined as being any numbered paragraph and all subparagraphs. Where different implementation responsibilities apply to different portions of a section, the implementation responsibility is specified on an individual

paragraph or sentence basis, as applicable. All implementing organizations designated shall accomplish their implementation activities in accordance with the requirements specified herein.



R. B. MacDonald
Manager, Large Area Crop Inventory Experiment

CONTENTS

Section		Page
1.0	FUNCTIONAL RESPONSIBILITIES	1-1
1.1	GENERAL	1-1
1.2	SPECIFIC	1-1
1.2.1	<u>Subsystem Implementation</u>	1-1
1.2.2	<u>Personnel</u>	1-1
1.2.3	<u>Procedures Development</u>	1-1
1.2.4	<u>System Operations</u>	1-1
2.0	<u>APPLICABLE DOCUMENTS</u>	2-1
3.0	<u>FUNCTIONAL FLOW DIAGRAMS</u>	3-1
3.1	SUBSYSTEM DEVELOPMENT	3-1
3.2	SUBSYSTEM OPERATION	3-1
3.2.1	<u>Plan</u>	3-1
3.2.2	<u>Implementation</u>	3-9
4.0	<u>CLASSIFICATION AND MENSURATION SUBSYSTEM REQUIREMENTS</u>	4-1
4.1	FULL-UP SYSTEM	4-1
4.1.1	Data Systems Requirements	4-1
4.1.2	<u>Training</u>	4-24
4.1.3	<u>Procedures Development</u>	4-25
4.1.4	<u>Green Number and Haze Correction Computations</u>	4-25
4.1.5	<u>Spectral Plots</u>	4-30
5.0	<u>SUBSYSTEM INPUT REQUIREMENTS</u>	5-1
5.1	FULL-UP SYSTEM	5-1
5.1.1	<u>DAPTS-Landsat Data Requirements</u>	5-1

Section		Page
5.1.2	<u>DAPTS-Field Data Requirements</u>	5-3
5.1.3	<u>DAPTS-Historical Agricultural Requirements</u>	5-4
5.1.4	<u>YES Requirements</u>	5-16
5.1.5	<u>CAS Requirements</u>	5-17
5.1.6	<u>ISRRS Requirements</u>	5-17
5.1.7	<u>Research Requirements</u>	5-21
5.1.8	<u>Test and Evaluation Requirements</u>	5-21
6.0	<u>CAMS OUTPUT REQUIREMENTS</u>	6-1
6.1	FULL-UP SYSTEMS	6-1
6.1.1	<u>Inputs to the ADP System</u>	6-1
6.1.2	<u>Data Status Reports</u>	6-2
6.1.3	<u>Classification Results</u>	6-3
6.1.4	<u>ADP Imagery</u>	6-7
6.1.5	<u>Outputs to DAPTS</u>	6-8
6.1.6	<u>Data Base Transaction Query Report</u>	6-9
6.1.7	<u>Resource Parameter Reports</u>	6-9
6.1.8	<u>Spectral Plots</u>	6-9
7.0	<u>INTERFACE REQUIREMENTS</u>	7-1
8.0	<u>OPERATIONAL REQUIREMENTS AFFECTING THE SYSTEM DESIGN</u>	8-1
8.1	THROUGHPUT REQUIREMENTS	8-1
8.2	RESPONSE REQUIREMENT	8-1
8.3	RELIABILITY REQUIREMENTS	8-1
8.4	SECURITY REQUIREMENTS	8-1
8.5	DELIVERY OF PRODUCTS	8-1
8.6	QUALITY ASSURANCE REQUIREMENTS	8-2

Section		Page
8.7	OTHER OPERATIONAL REQUIREMENTS	8-2
9.0	<u>CAMS VERIFICATION REQUIREMENTS</u>	9-1
9.1	DATA SYSTEMS VERIFICATION	9-1
9.1.1	<u>Design Reviews</u>	9-1
9.1.2	<u>Verification Testing</u>	9-1
9.2	ANALYST CERTIFICATION	9-1
9.3	PROCEDURES VERIFICATION	9-1
9.3.1	<u>Operational Procedures</u>	9-1
9.3.2	<u>Analysis Procedures</u>	9-1
10.0	<u>RESEARCH REQUIREMENTS</u>	10-1
11.0	<u>TEST AND EVALUATION</u>	11-1
APPENDIX		
A	REQUIREMENTS FOR CAMS INTERFACE REPORT	A-1

FIGURES

Figure		Page
3-1	CAMS development flow	3-2
3-2	Functions of the classification and mensuration subsystem (CAMS)	3-3
3-3	Illustration of local signature extension	3-4
3-4	Illustration of analog site signature extension	3-5
3-5	Signature extension procedure	3-8
3-6	Training segment selection	3-10
3-7	Overview of CAMS processing	3-11
3-8	Functional overview of LACIE CAMS	3-12
3-9	LACIE CAMS functional image flow	3-13
3-10	LACIE CAMS field definition flow	3-15
3-11	LACIE classification	3-16
5-1	United States ancillary summary	5-7
5-2	Soviet Union ancillary summary	5-9
5-3	Crop calendars	5-12

GLOSSARY

ADP	Automatic Data Processing
AI	Analysis Interpretation (of Imagery) or Analyst Interpreter
C&I	Composition and Indexing
CAMS	Classification and Mensuration Subsystem
CAS	Crop Assessment Subsystem
CCBD	Change Control Board Directive
CCT	Computer Compatible Tape
CITARS	Crop Identification Technology Assessment for Remote Sensing
CMS	Classification and Mensuration Subsystem (Obsolete)
CPU	Central Processing Unit
DAPTS	Data Acquisition Preprocessing and Transmission Subsystem
DAS	Data Analysis Station
DO	Designated Other
DOY	Day of Year
DPA	Data Processing Analyst
DR	Discrepancy Report
DU	Designated Unidentifiable
EOD	Earth Observations Division
EOF	End of File
ERIM	Environmental Research Institute of Michigan
ERIPS	Earth Resources Interactive Processing System
ERTS	Earth Resources Technology Satellite (Obsolete, see Landsat)
GSFC	Goddard Space Flight Center
IA	Image Analysis

ICD	Interface Control Document
IE	Information Evaluation
I/O	Input/Output
ISRRS	Information Storage, Retrieval, and Reformatting Subsystem
JSC	Johnson Space Center
LACIE	Large Area Crop Inventory Experiment
LACIP	Large Area Crop Inventory Project (Obsolete)
Landsat	Landsat (Formerly Earth Resources Technology Satellite, ERTS)
LEC	Lockheed Electronics Company
MLA	Mean Level Adjustment
MSS	Multispectral Scanner
MTU	Magnetic Tape Unit
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration
ONC	Operational Navigation Charts
PI	Photo Interpreter, (See AI)
PFC	Production Film Converter
PTD	Photographic Technology Division
R&D	Research and Development
RTEB	Research, Test, and Evaluation Branch
SPE	System Performance Evaluation
TBD	To Be Determined
TLU	Table Look Up
USAF	United States Air Force
YES	Yield Estimation Subsystem

SECTION 1.0
FUNCTIONAL RESPONSIBILITIES

- 1.1 GENERAL
Design, implement, and operate a classification subsystem to determine values for the proportion of wheat in each Landsat acquired sample unit defined by the Crop Assessment System (CAS), at each of as many as four periods in the growth of wheat. The subsystem shall be designed to utilize Landsat data as the primary source.
- 1.2 SPECIFIC
 - 1.2.1 Subsystem Implementation
Develop, specify, and implement the hardware and software systems required to meet the operational requirements of section 1.2.4.
 - 1.2.2 Personnel
Identify and train the personnel needed to operate the above system in accordance with the procedures of section 1.2.3.
 - 1.2.3 Procedures Development
Develop procedures to accomplish the operational responsibilities of section 1.2.4 with the system of section 1.2.1 and the personnel of section 1.2.2.
 - 1.2.4 System Operations
Estimate, for each acquisition of each sample unit, the proportion planted in wheat and, for those acquisitions occurring after wheat harvest, the proportion of wheat which was harvested. Estimate the error in each of the above estimates. Determine the resources consumed in the above operations. Report all of the above to the CAS. Support CAMS Post Operation Performance Analysis by Testing and Evaluation. Perform additional analyses of sample segments in support of accuracy assessment and Test and Evaluation.

SECTION 2.0
APPLICABLE DOCUMENTS

The following documents are applicable to the extent specified within:

- A. Large Area Crop Inventory Experiment Project Plan - LACIE-C00605, August 1975.
- B. Earth Resources Interactive Processing System Requirements Document PHO-TR514, June 1974
- C. Flowcharts for LACIE Operational CAMS, December 2, 1974
- D. LACIE CAMS Training Plan LACIE-00601, April 1975

REPLACING PAGE BLANK NOT FURNISHED

SECTION 3.0
FUNCTIONAL FLOW DIAGRAMS

3.1 SUBSYSTEM DEVELOPMENT

Figure 3-1 indicates the task flow required to place the CAMS into operation.

3.2 SUBSYSTEM OPERATION

Figure 3-2 shows the basic CAMS functions in terms of gross inputs and outputs.

3.2.1 Plan

Operation of the CAMS is based upon multispectral/multitemporal pattern recognition of Landsat data. In this process, the identity of the ground cover class for each Landsat picture element ("pixel") is estimated by comparing its radiance in four spectral bands for each of as many as four dates to the distributions of radiance (versus spectral band and date) for various candidate ground cover classes. In this classification process, the distributions are assumed to be multivariate normal, and the decision as to the most likely class is made in a maximum likelihood fashion.

Such a classifier requires statistics (mean vectors and covariance matrices) which describe the radiance distributions for the various candidate classes. In the CAMS, such statistics are estimated by computing the means and covariance matrices from actual Landsat data over earth surface areas covered by the actual ground cover classes of interest.

These distributions ("signatures") are functions of many variables such as soil color and moisture content, growth stage of plants, atmospheric haze and water content, and solar elevation angle; therefore, in the CAMS, the statistics must be computed for regions in which these parameters are similar to those expected in the region to be classified. Additionally, the CAMS provides capabilities of correcting for changes in solar elevation angle and atmospheric effects.

Two approaches will be used to find suitable areas in which to train the classifier prior to classifying a particular segment of Landsat data. In the first approach, analysis interpretation (AI) of imagery produced from Landsat segment data will be performed to find areas of wheat and non-wheat ground cover classes. Statistics will be computed over these areas ("training fields") and used to classify other pixels in the same Landsat segment or in other nearby segments. In the second approach, "analog regions" in which the signatures of wheat and non-wheat are highly similar to those in the recognition segment will be located. Then, signatures will be computed for areas of known ground cover classes and extended to the desired recognition area. These processes are indicated in figure 3-3 and 3-4.

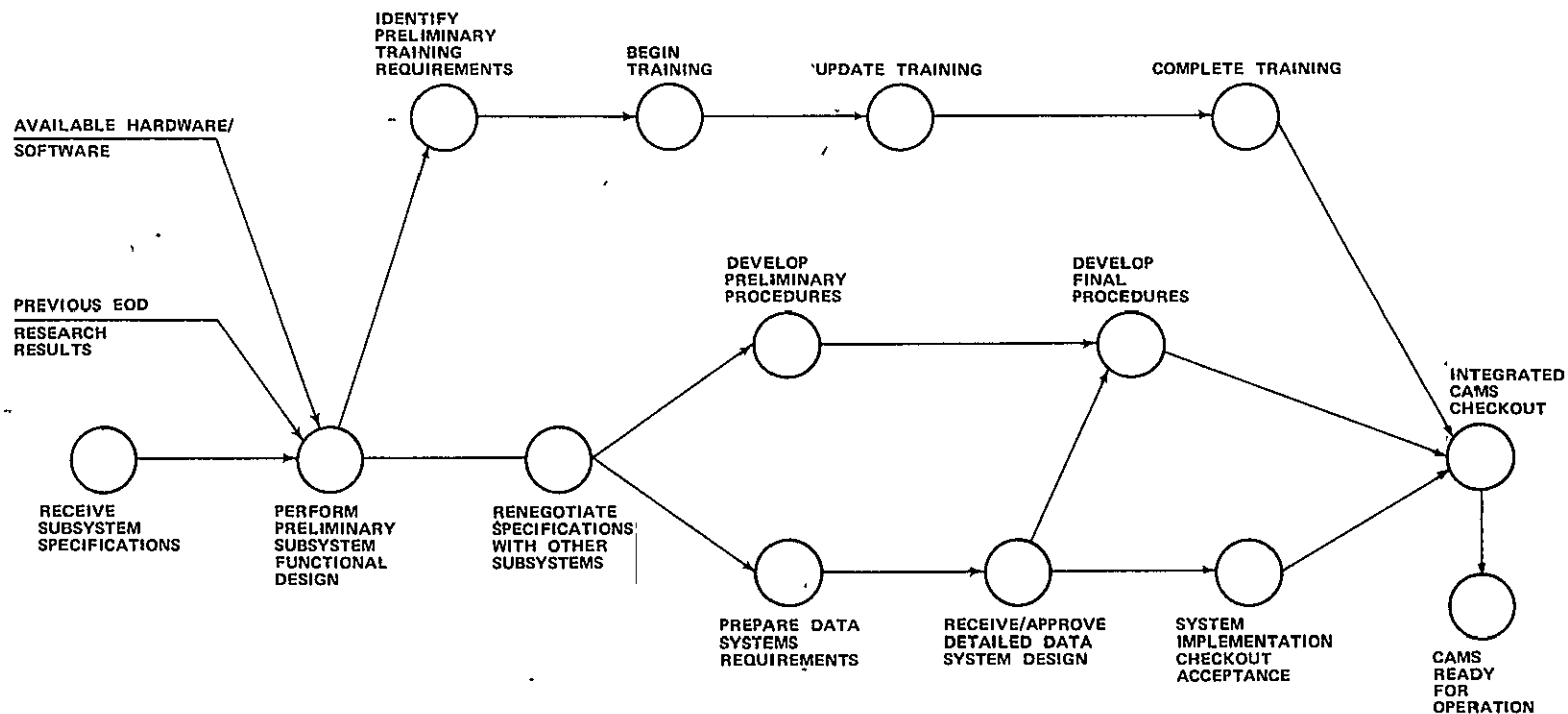


Figure 3-1.- CAMS development flow.

ORIGINAL PAGE IS
OF POOR QUALITY

C A M S F U N C T I O N S

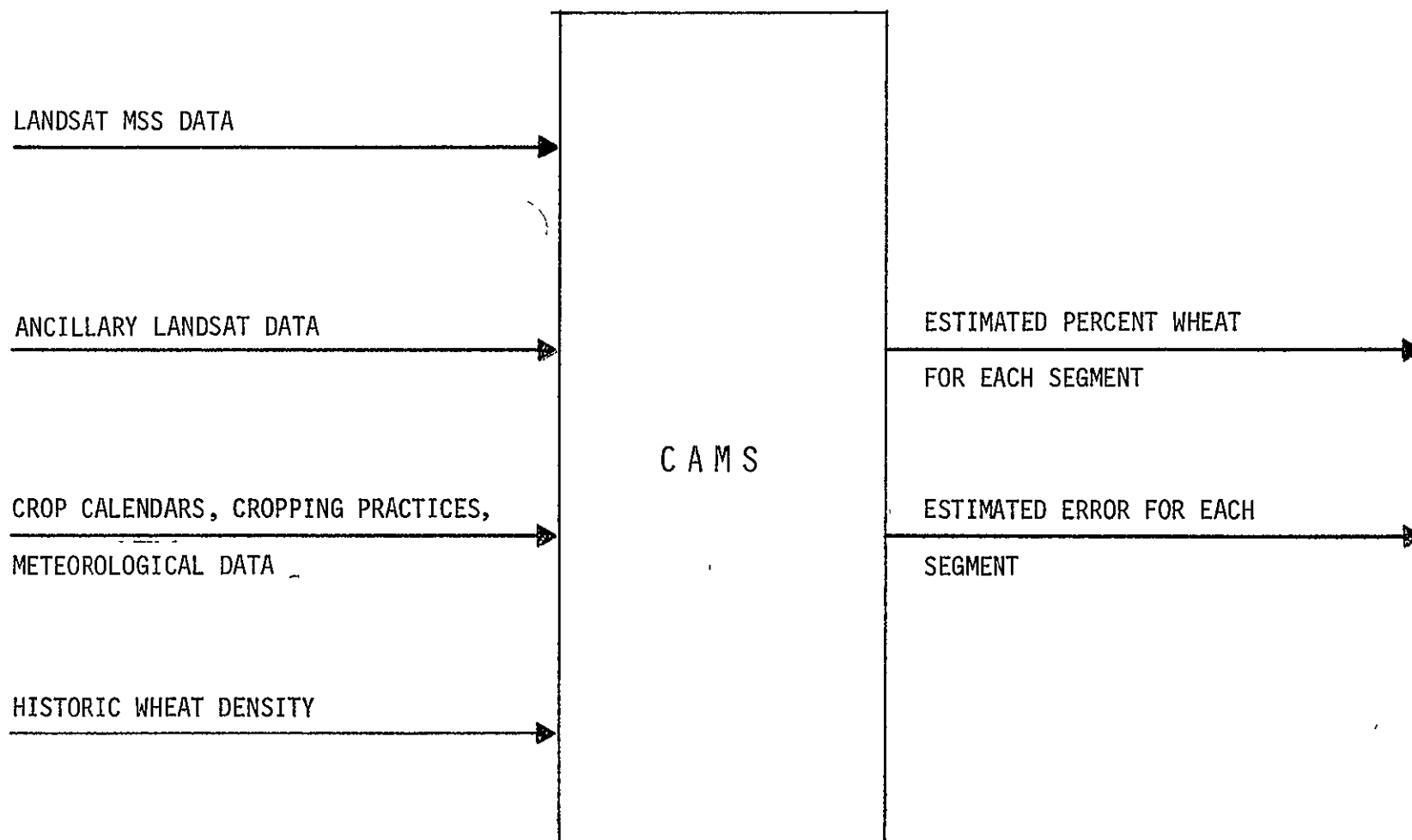


Figure 3-2.- Functions of the classification and mensuration subsystem (CAMS).

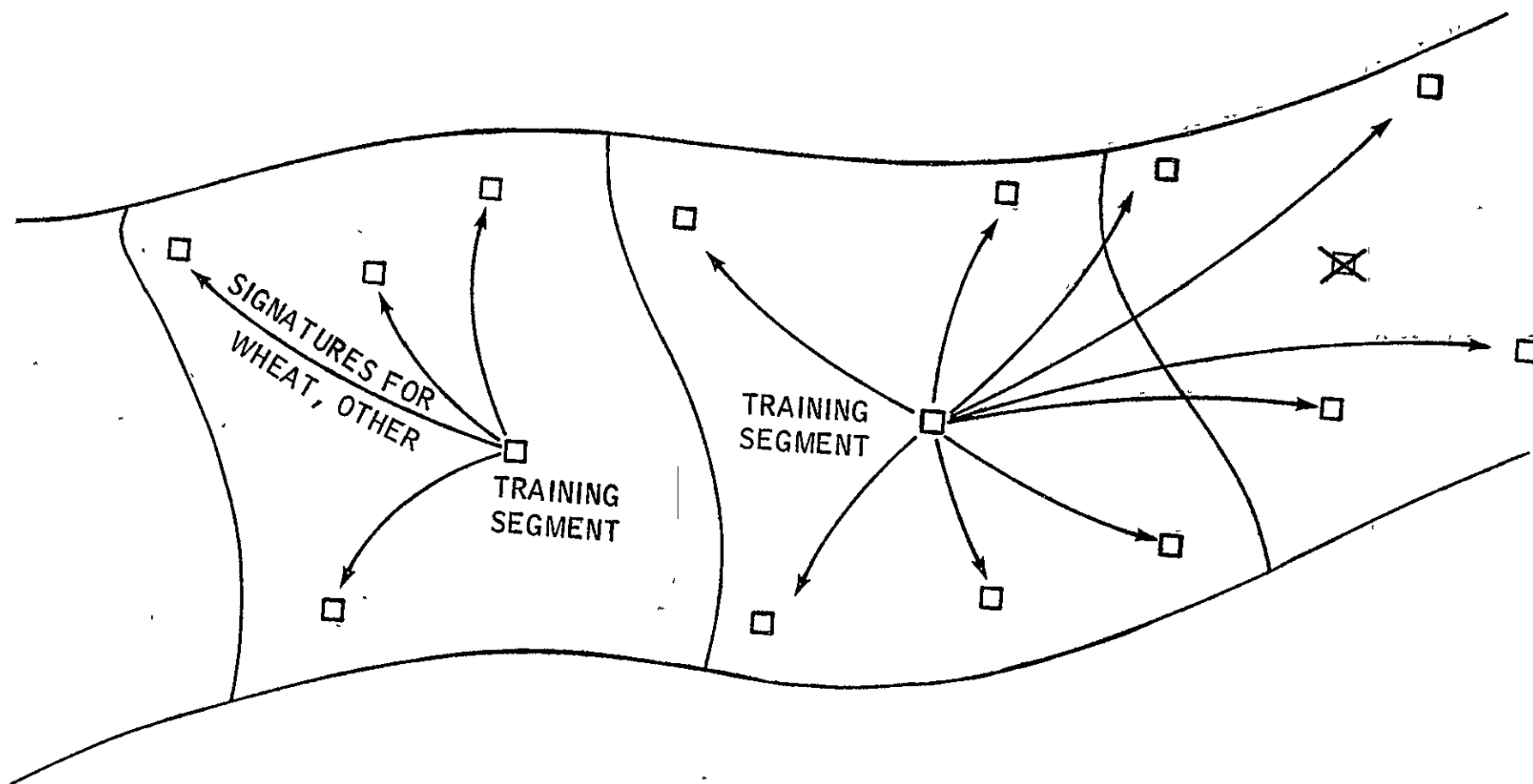


Figure 3-3.- Illustration of local signature extension.

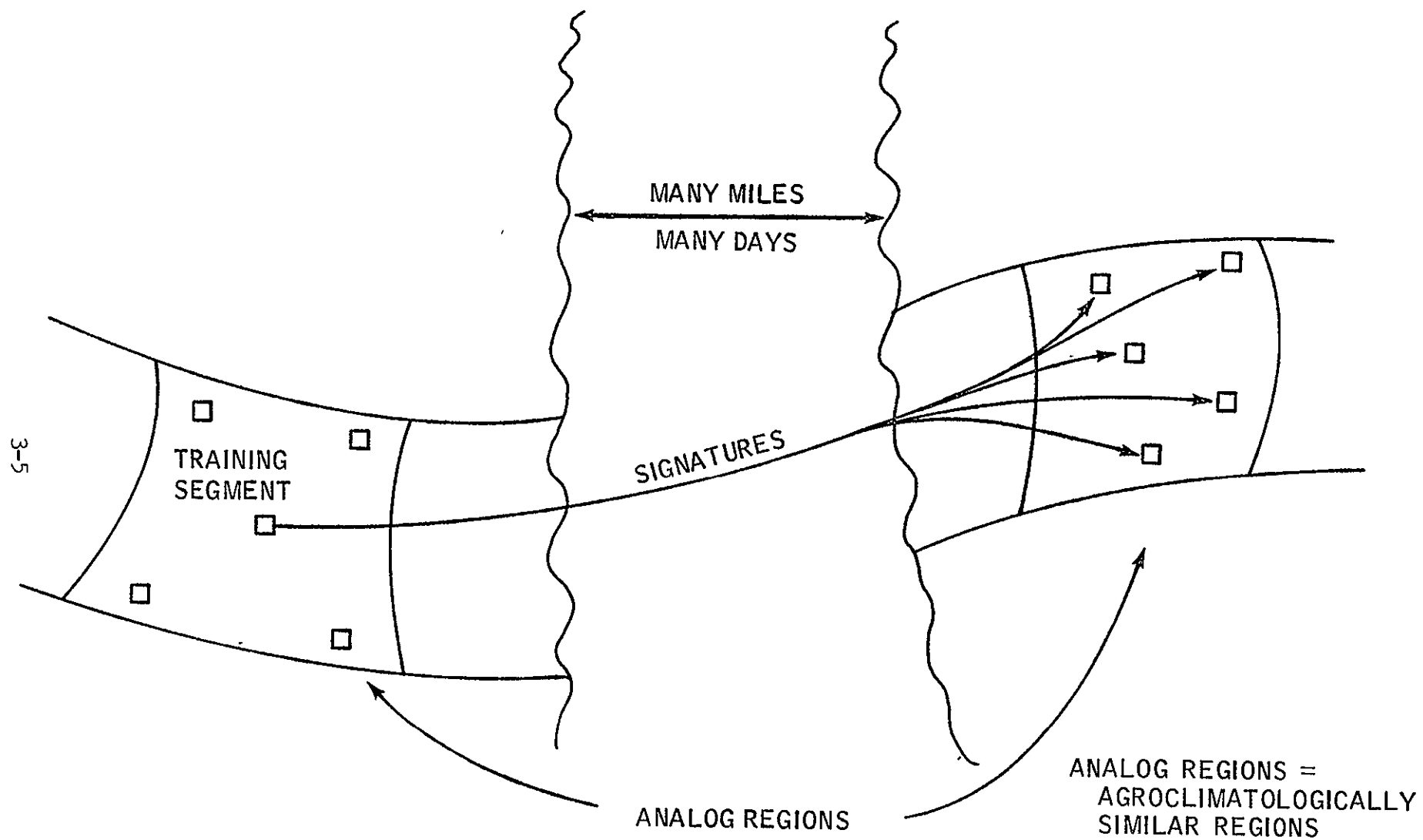


Figure 3-4.- Illustration of analog site signature extension.

Image analysis to locate training fields will not be performed in all segments. It is more cost effective to locate training fields in an effective subset of the segments and extend the signatures to all segments. It is estimated that this subset of the segments ("training segments") will contain 20 percent of all segments.

Image analysis will also be used for another purpose. Previously, difficulty has been experienced in accurately classifying such non-cropland areas as forests, and towns. Frequently, such areas are highly inhomogeneous at the Landsat resolution, and poor classification has resulted. Therefore, in the CAMS, major noncropland areas will be identified by image analysis and will be manually excluded. Similarly, areas in which data quality is poor due to such factors as clouds and noise can be identified manually.

Wheat-growing regions of the earth will be partitioned into smaller areas over which signature extension is expected to be successful. This partitioning will be accomplished based upon such ancillary data as crop calendars, meteorology, and soil color, as well as on the basis of trial classifications. Each signature extension region will contain one or more training segments. If cloud cover or other operational problems prevent acquisition of the training segment for a given signature extension region, signature extension will be attempted from a neighboring region. This approach is expected to provide reduced, but still acceptable, classification accuracies in the region without a training segment. It is anticipated that signature extension regions typically will have approximately the area of a Landsat frame. This approach is also indicated in figure 3-3.

The CAMS design is based upon the estimate that wheat can be separated adequately from other crops by analysis of up to four acquisitions of Landsat data during the growth of wheat. The Landsat acquisitions are to occur in a sequence which maximizes the discrimination of wheat from non-wheat and the usability of the obtained information.

Nominal Biowindow Definitions (Historical Data)

A. Biowindow I

1. Winter Wheat (South of 44° N lat) - U.S., U.S.S.R.
Opens one month past when 50 percent of the strata was planted or at stage 2.3 (when crop is 30 percent between emergence and jointing). In other wording, stage 2.3 is estimated as 30 days after 50 percent planting date. Close is one day before 50 percent of the strata has been reported in the jointing stage.
2. Winter Wheat (North of 44° N lat) - U.S., U.S.S.R.
Open same as above. Closes when snow cover dominates the image or when climatological average one inch snow line occurs or about 50 days after stage 2.3.

B. Biowindow IB

Winter Wheat (North of 44° N lat) - U.S., U.S.S.R.
Opens 50 days before 50 percent jointing (based on nominal SRS reporting data for 50 percent jointing). Closes one day before 50 percent of the strata has been reported in the jointing stage.

C. Biowindow I

Winter Wheat - Brazil, Argentina, Australia, India, China
Opens at the 50 percent planted date. Closes one day before 50 percent of the strata has been reported in the jointing stage.

D. Biowindow I

Spring Wheat
Opens 15 days before 50 percent of the strata has been reported in the planting stage. Closes one day before 50 percent of the strata has been reported in the jointing stage.

E. Biowindow II

Winter Wheat and Spring Wheat
Opens when 50 percent of the strata has been reported in the jointing stage. Closes one day before 50 percent of the strata area has been reported in the heading stage.

F. Biowindow III

Winter Wheat and Spring Wheat
Opens when 50 percent of the strata has been reported in the heading stage. Closes one day before 50 percent of the strata has been reported in the soft dough or turning-yellow stage.

G. Biowindow IV

Winter Wheat and Spring Wheat
Opens when 50 percent of the strata has been reported in the soft dough or turning-yellow stage. Closes 18 days after 50 percent of the strata area has been reported harvested.

For LACIE Phase III, the above also is acceptable for all training segments.

For LACIE Phase III, window IB will be used only for those segments for which no satisfactory interpretations were made in window I.

All segments will be acquired by GSFC at every available opportunity; however, to reduce the cost of preprocessing, transmitting, and storing the data, ordinary segments will be preprocessed and transmitted to JSC only once (if acquired) per stage. Designated training segments will be preprocessed and sent to JSC for each acquisition lying within a development stage. This is indicated

DEVELOPMENT STAGE	ACQUISITION	SEGMENT #	1	2	3	4	5	6	7	8	9	10
CROP ESTABLISHMENT	1		X	⊗	X	X			X			
	2			⊗			X	X	X		X	X
	3											
GREEN	1				X			X	⊗	X	X	
	2		X	⊗		X	X					
	3			⊗								X
HEADING	1		X	⊗	X	X	X	X	⊗	X	X	X
MATURE	1		X	⊗				X		X		X
	2			⊗	X	X	X		⊗		X	

Figure 3-5.- Signature extension procedure.

in figure 3-5 in which acquisitions of ordinary segments are indicated by X's and acquisitions of designated training segments are indicated by circled X's. The purpose of the extra acquisitions is explained below.

Signatures obtained on one acquisition within a development stage are not necessarily expected to be valid for other acquisitions within the same phase; however, training fields obtained on one acquisition will usually be valid on other acquisitions within the same phase. Therefore, by reacquiring the training segments and using training fields located on previous acquisitions, statistics appropriate to an acquisition can be computed without expending image analysis on that acquisition. It is planned to conduct major image analyses only upon training segments and only once per development stage.

Figure 3-5 also indicates, in the column for segment 7, that where no suitable training data are acquired, it will be necessary to perform additional image analysis. Such analysis will include not only the most recent acquisition but also previous acquisitions. Changes in training segment selection will be made under procedural control.

Selection of training segments for use in classifying individual recognition segments will be a dynamic process. Figure 3-6 indicates the inputs, considerations, and outputs in this process.

3.2.2

Implementation

Figure 3-7 gives a superficial view of the data processing required to implement the foregoing plan. Incoming Landsat data (which arrive at JSC one acquisition at a time, but are preregistered so that ground positions on different acquisitions bear the same scanner coordinates) are used to generate film for image analysis and are input to the computer for composition and indexing (C&I). The C&I application stores the newly acquired data and prepares indexed data reports. (The C&I application is not discussed further in this volume, as it is a part of the Information Storage, Retrieval, and Reformatting Subsystem (ISRRS).) Working with the film and the indexed data reports, analysts enter field coordinates and job control input to the computer for classification processing. Completed classification jobs are screened for classification quality, and, where necessary, jobs are reworked.

Figure 3-8 shows that the classification processing will be operated in two modes, a semiautomatic (batch) mode and an interactive mode. The major function of the data processing is expected to occur in the batch mode, with the interactive mode being reserved for those jobs not practicable in the batch mode.

Figure 3-9 shows a more detailed view of the image generation process. Incoming Landsat data are composed with previously acquired Landsat data, allowing the generation of multitemporal

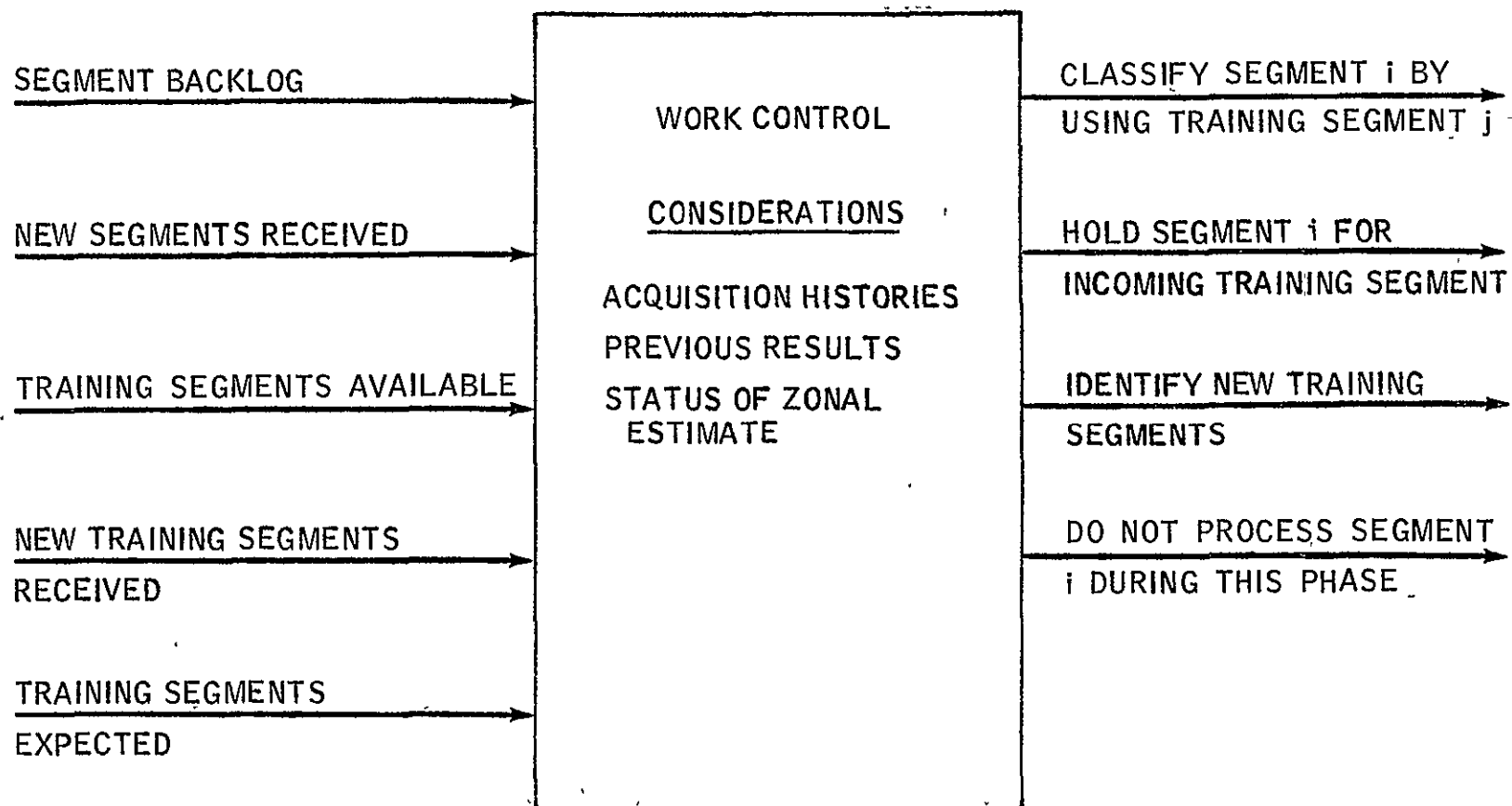


Figure 3-6.- Training segment selection.

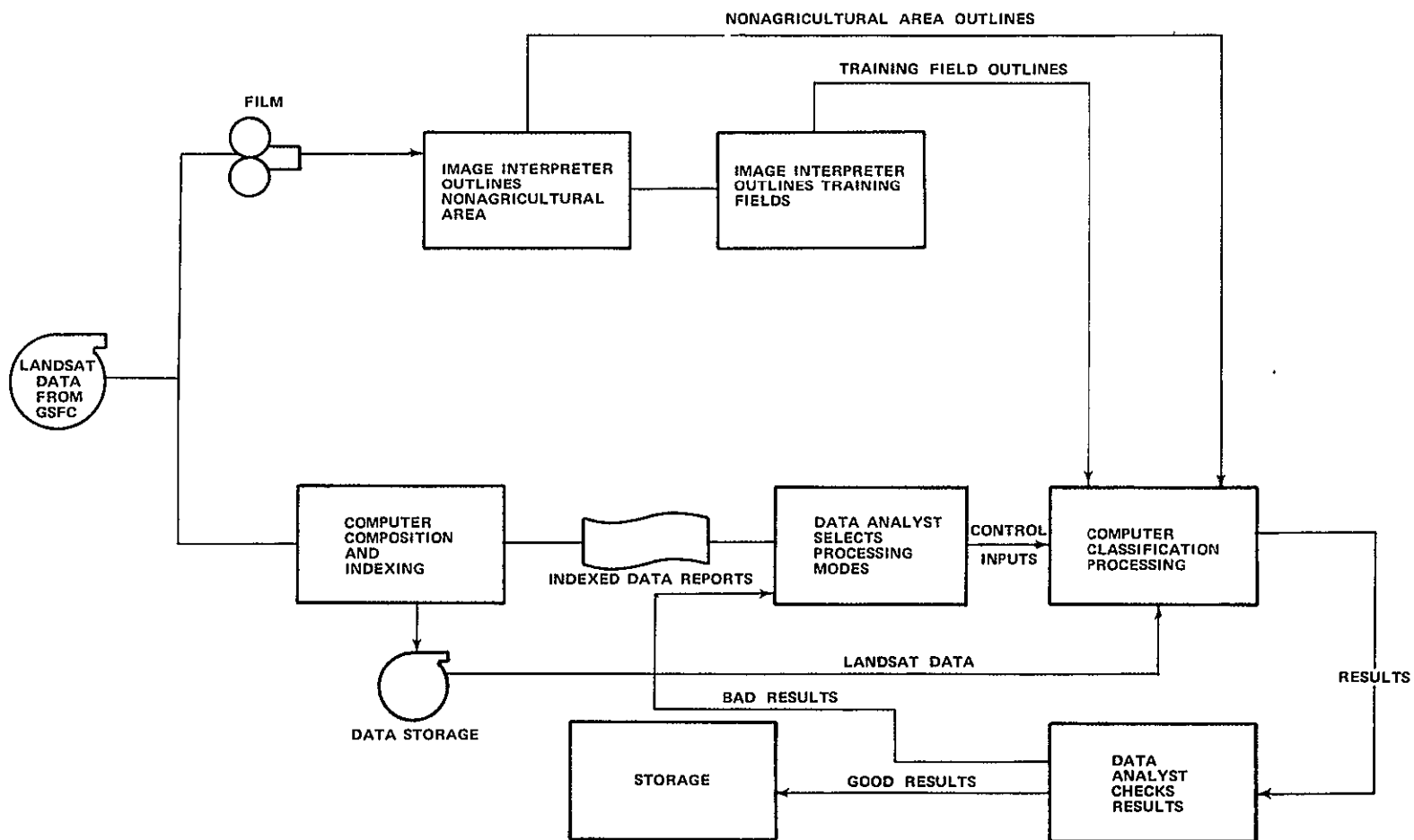


Figure 3-7.- Overview of CAMS processing.

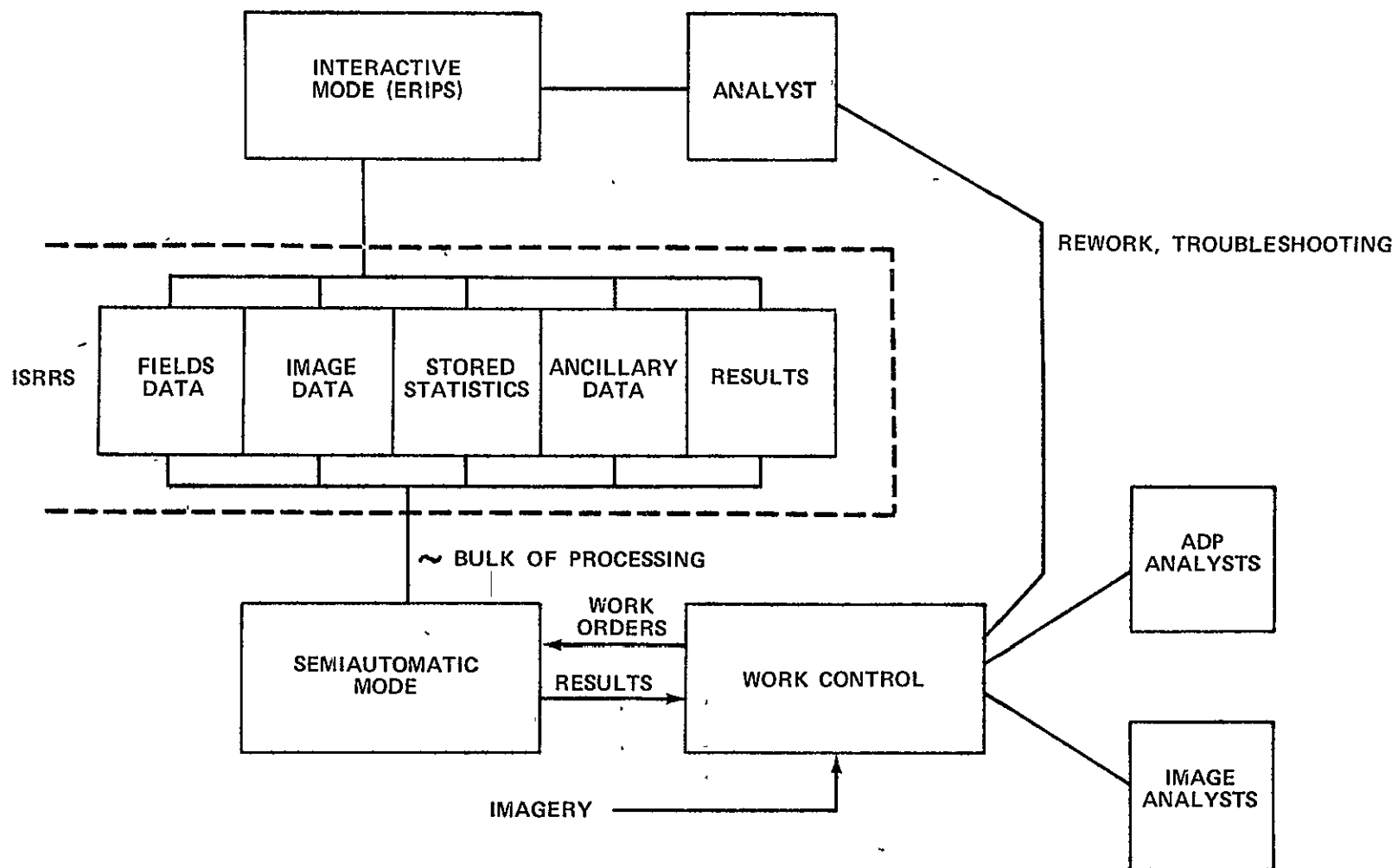


Figure 3-8.- Functional overview of LACIE CAMS.

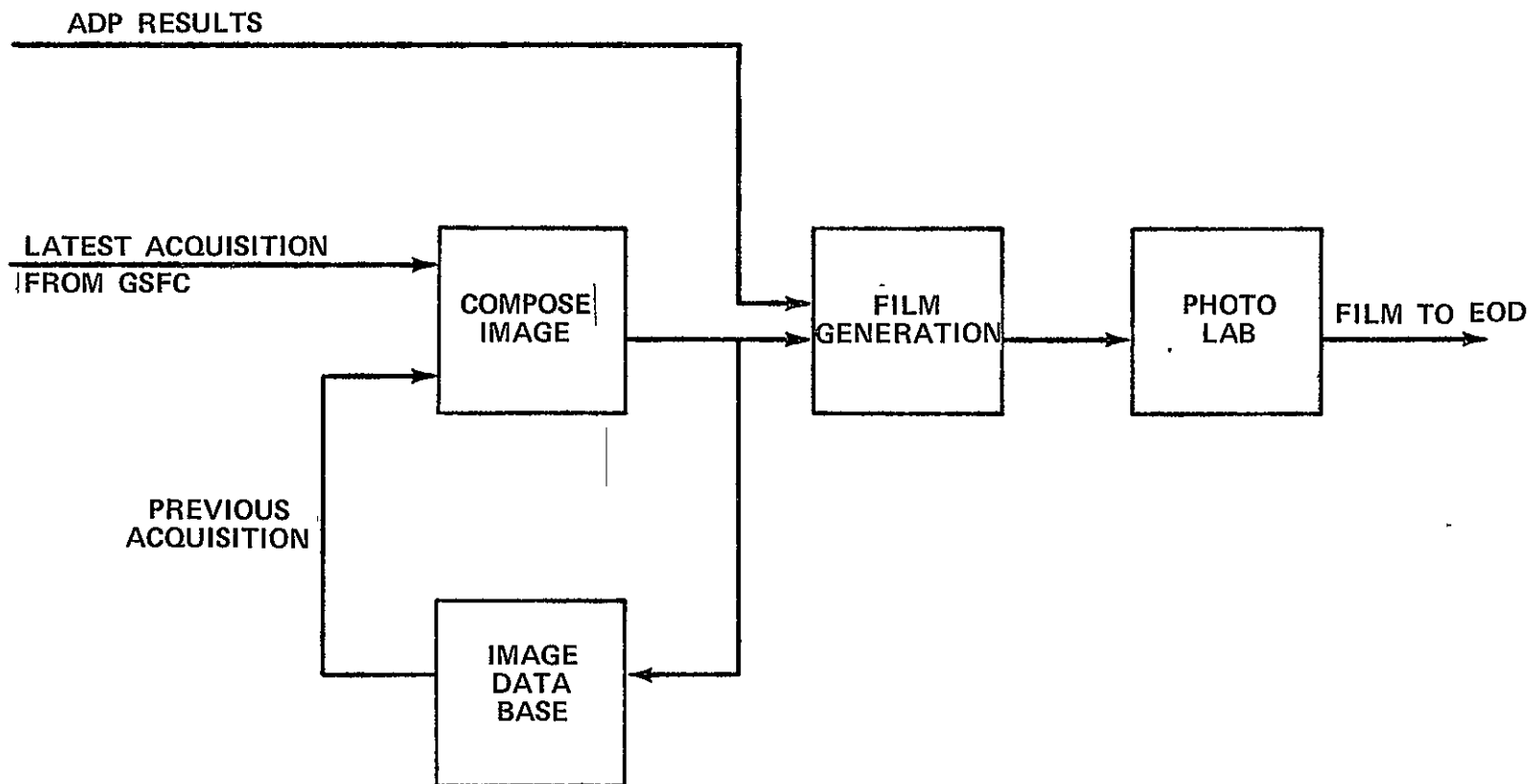


Figure 3-9.- LACIE CAMS functional image flow.

images. Additionally, the composed data are returned to the image data base where they are available for use in classification processing. Film is generated using the Production Film Converter (PFC) in building 30, or as a backup, the Aerojet Data Analysis Station (DAS) in building 17. Film is transmitted to the JSC Photographic Technology Division (PTD) for development and then delivered to the Earth Observations Division (EOD) for analysis. Note that some ADP results are also transmitted in the form of imagery.

Figure 3-10 shows the flow of field definitions obtained by Analyst Interpreters (AI) and Data Processing Analysts (DPA). Field definitions, in scanner coordinates, are input, via a work control center, to the fields data base in the ISRRS (formerly IMS). The ISRRS issues, via the PFC and the PTD (not shown), a transparency at the same scale as other LACIE imagery showing the boundaries and identities of all fields.

Figure 3-11 is a more detailed view of the batch mode classifier. The interactive classifier is similar, except that the analyst, working at a terminal, replaces the CAMS supervisor. In the classifier, training field coordinates, image data, and instructions are input to the statistics computation process. Following statistics computation, modifications to statistics to account for changes in sun angle or atmospheric turbidity can be performed. Computed statistics are available as an output at this point; alternatively, external statistics can be input at this point. This allows the development of offline statistics modification programs. The statistics can then optionally be input to the feature extraction processor in which those Landsat channels containing the most information can be extracted. Next, image data from the recognition segment and statistics are input to the classifier, which classifies each pixel in one of the input categories. The classifier offers the option of summing the density functions of a number of classes and/or subclasses into a category density function.

Following classification, certain critical results are input to the ISRRS where they are accessible to the CAMS and other subsystems. More extensive results, including diagnostics and all data required to duplicate the run, are output to the analyst. Following evaluation of the results, the analyst stores an evaluation of the results in the ISRRS.

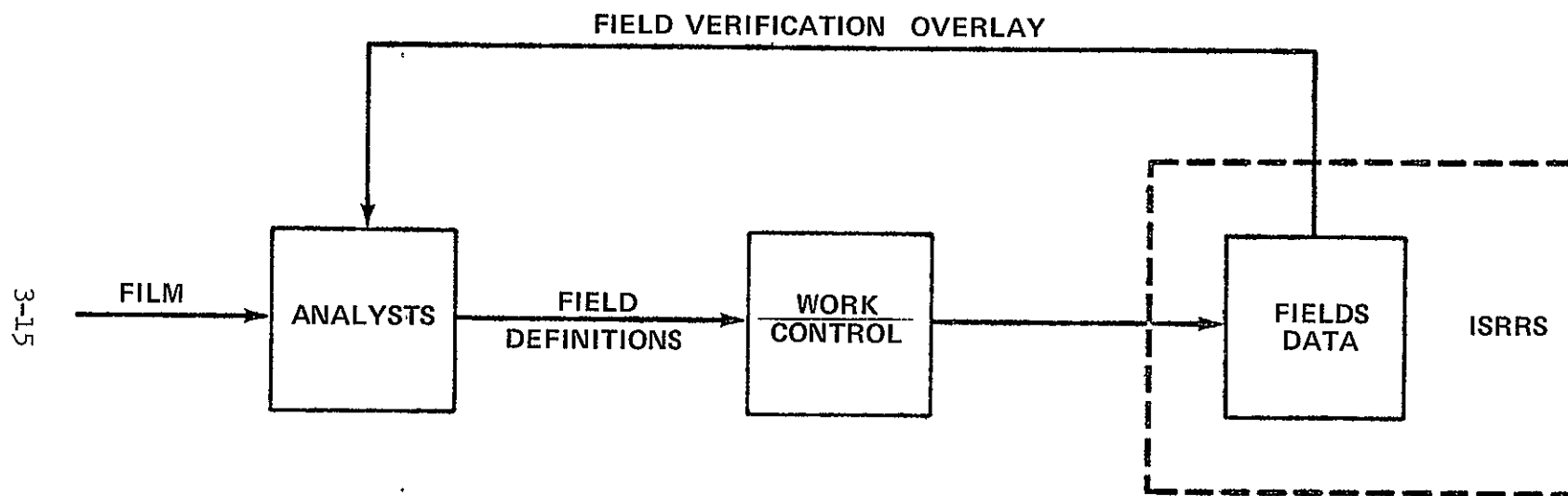


Figure 3-10.- LACIE CAMS field definition flow.

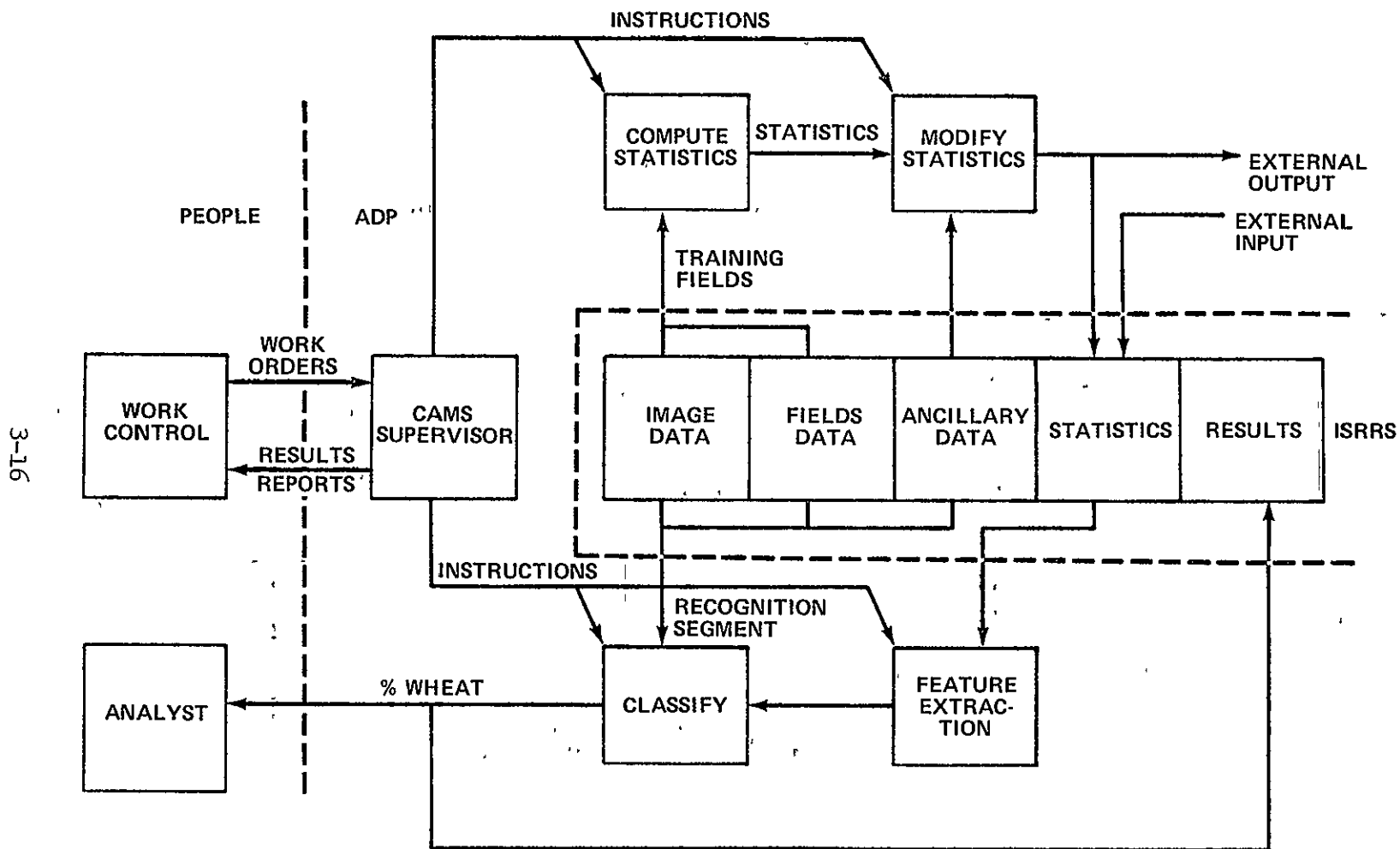


Figure 3-11.- LACIE classification.

SECTION 4.0
CLASSIFICATION AND MENSURATION SUBSYSTEM REQUIREMENTS

4.1 FULL-UP SYSTEM

4.1.1 Data Systems Requirements

(Req'd by CAMS; Cat. 1, except as noted; Impl. Resp.: CAMS/DSAD)

This subsystem is envisioned as several applications programs capable of being operated in several modes. The design is required to be modular so as to allow subsequent additions and changes in algorithms and procedures with minimum delay.

The system shall be a production line system consisting of a relatively inflexible semiautomatic mode and a highly flexible interactive mode, both of which are of operating on a common data base. This data base will contain extensive Landsat segment data. While it is not required that this entire imagery data base be simultaneously accessible, it is required that some means be provided to ensure that those segments to be processed together be simultaneously accessible when needed.

The system must be designed so that the following conditions are met:

- A. An average throughput of (420 Phase I, 714 Phase II, and 840 Phase III) sample segments/month can be processed, after allowance for downtime.
- B. Turnaround time on each processing job is less than 24 hours.
- C. Any segment can be processed on any day, regardless of the processing status of other segments.

Status reports indicating the status of all Landsat segment data in the system are required. It shall be the responsibility of the EOD to verify performance of processing jobs.

The semiautomatic mode should be designed to minimize manual handling, particularly transcription, of processing instructions.

This Requirements Document references flow diagrams which are intended to clarify the verbal description of the processing. While these flow charts necessarily imply an implementation, the implementation shown should be considered only illustrative. Any other implementation which is functionally equivalent is clearly acceptable.

Flexibility shall be provided to allow change of control parameters in the semiautomatic mode under configuration control. This application program must be able to perform classification on any number of categories up to 26. Each category may be divided into one to 60 classes. Each class may be subdivided into one to 60 subclasses. The total number of subclasses for all categories combined shall not exceed 60.

- 4.1.1.1 Channel requirements.-- Purging and retention of these data are TBD, but current planning should be based upon the assumptions that image data will be retained for active use three months after the end of a crop year and that purging and reinitiation will occur upon a crop year basis.

During this TBD period, data from two crop years may exist in the active image data base for some segments. The capability to purge either year is required. This option must be available for Phase II.

Channels and acquisition dates to be used in further processing must be selectable on a per segment, by user input.

- 4.1.1.2 Fundamental unit.-- The fundamental unit of data to be processed at each distinct execution of this application program shall be a sample segment or a collection of sample segments to be classified, and a set of one to three training segments. Member(s) of the group of sample segment(s) to be classified may also be members of the group of training segments.

- 4.1.1.3 Fields.-- The user shall be able to identify the following types of fields in each segment:

- A. Class training fields (up to 60 subclasses)
- B. Class test fields (up to 60 subclasses)
- C. Designated Other (DO) fields
- D. Designated Unidentifiable (DU) fields
- E. Mean Level Adjustment (MLA) fields

- 4.1.1.3.1 Data base:

- A. A fields data base, which may be vacant, shall exist for every segment, and contain up to 146 fields per each segment.
- B. A dot data base shall exist which will contain a set of fixed single pixel locations, with the category level designation and function designation for each segment. The data base is updated by an offline utility program.
- C. The capability should exist to store multiple sets of dots and fields for each segment so that any set of dots and the corresponding set of fields can be used to process a segment.

When a segment is represented in the active image data base by data from two crop years, separate fields data bases shall exist for the two separate years.

These fields shall be used as follows:

- A. Training fields will be used as the initial source of class or subclass statistics.
- B. Test and/or training fields will be used to evaluate classifier performance.
- C. DO fields will automatically be identified as the category "Other," regardless of pixel data values, except when processing test fields, or training fields for the category "Other."

- D. Pixels lying in DU fields will automatically be identified as belonging to the category Unidentifiable, regardless of the pixel data values, with the following two exceptions. Pixels lying in both DO and DU fields will be identified as belonging to the category "Other." When classifying test fields, or training fields for the category "Unidentifiable", all pixels will be classified on the basis of pixel data values.
- E. MLA fields will be used in an algorithm (below) to compensate for atmospheric variations.

4.1.1.3.2 Field definition: The system shall generate a field verification overlay at the same scale as LACIE AI imagery.

Field definitions to be used in further processing shall be obtainable by each of the following routes:

- A. 1. By input of coordinates selected externally to the system.
2. By output and reinput of coordinates selected interactively on the system.
- B. Interactively by online field definition. In this mode, the analyst may request recall of the stored coordinates as a starting point. The interactive operation, however, will not update the fields data base.

Fields can contain any number of pixels, pixels/line, or lines per column from one up to the limit set by segment dimensions. Fields may have from 2 to 10 vertices. Minimum field width shall be one pixel.

4.1.1.3.3 A 'replace' option should exist for the Fields Data Base. The 'replace' option would allow existing fields and statistics to be deleted from one category, and in a single operation, then be included in a different category.

4.1.1.3.4 The system shall generate an overlay of current dot labels at the same scale as LACIE AI Imagery. The overlays shall be on the same scale as the PFC Product 1. The overlay shall be transparent with each dot designated by a unique color or gray shade by category. This color or gray shade shall be consistent between segments. The overlay should carry the data of the update and the segment identification.

Interactively, dots, labels and types to be used in further processing shall be obtainable by each of the following routes:

- A. By input of coordinates, labels and types selected externally to the system.
- B. By online dot definition. Analyst may request recall of stored dots as starting point.
- C. Dots with labels and types shall be displayed as an overlay to the television monitor, either color or gray level. Present labels are W, N, X, DO, and DU. Dot types are S for starting vector, L for labeling vector or B for bias correction vector.

- D. Dot definition changes shall be controlled by analyst via keyboard inputs. Modification to the dot data base by analyst should be restricted by use of read/write keys to insure data base security.

4.1.1.4 Computation of statistics.- The system shall be capable of aggregating the training fields from up to three segments for all further processing defined in this section. That is, the system shall be able to process these fields as though they were located in a single segment. The system shall also be capable of merging statistics obtained in up to three nonaggregated segments.

Three modes shall be available for computing statistics from training data as follows:

Mode 1: Input training fields are used to directly define class or subclass statistics without further operations. This mode is the default option.

Mode 2a: Input training fields are used to provide starting vectors and data points for clustering algorithm. Clustering algorithm develops clusters. Cluster chaining algorithm is used to chain similar clusters together. The (chained) clusters are then used to compute subclass statistics.

Mode 2b: As above without chaining of clusters.

Hard copy output products from the clustering modes will consist of the following:

A. Cluster Map: Two types of color coded cluster maps shall be produced in transparency form at the same scale as other segment imagery (ref. sec. 5.1.6.1) whenever a segment is batch processed. These cluster maps shall provide the capability to display all the clusters produced in any single clustering run.

A.1 An unconditional cluster map shall be produced in which the pixels belonging to each cluster are displayed as a different color where the colors are chosen in order from a color table. A capability shall be provided to input a default color table.

A.2 A conditional cluster map shall be produced in which all unconditional clusters with the same category label are displayed as the same color but with a different color for each category. In addition, each conditional cluster shall be displayed using a different color. A conditional cluster is a cluster whose L1 distance from the subclass or dot used to label it to the cluster mean is greater than a user input threshold. The colors used to produce this map shall be selected in order from a color table. A capability will exist to input a default color table.

In order to aid the analyst in reviewing the cluster maps produced by a batch job, each map shall be annotated with a color key in the form of an example of each color that was used to make the map in the order in which the colors appear in the color table. The color key for a cluster map shall appear adjacent to the map i.e. between the image and the header (ref. sec. 5.1.6.2.6 - Image Header).

Cluster maps shall indicate the boundaries and identities of all fields by a technique which does not prevent the user from ascertaining the cluster of any pixels. It is permissible to use multiple sheets to accomplish this.

- B. Reports comparable to the ERIPS detailed clustering reports including options used in clustering, cluster-class matches, cluster population and cluster names.
- C. Distance tables including intra- and interclass distances, for all classes included in the clustering run.
- D. A report for each cluster generated which shows categories assigned to that cluster, the labeling dot closest to that cluster, and a listing of individual dots assigned to the cluster with the distance from each dot to the cluster mean. The report shall contain the information in the following format:

Cluster Name ____; Labeling Dot ____; Categories ____;

<u>DOTS:</u>	<u>DISTANCE:</u>	<u>DOTS:</u>	<u>DISTANCE:</u>
_____	XX,X	_____	XX.X
_____	'	_____	'
_____	'	_____	'
_____	'	_____	'
_____	XX.X	_____	XX.X

Mode 3: Input training fields are assumed to represent clusters. The Bhattacharyya distance measure is used to decide whether or not to chain fields together. The (chained) fields are then used to compute subclass statistics.

Hard copy output products from Mode 3 shall consist of the following (menu to be furnished):

- A. Intraclass distance tables $H[(i,j), (i,k)]$ for all fields in each class (see flow charts for details).
- B. List of chained fields. Number of pixels per chain.

The modes shall be individually selectable by class by segment.

The clustering algorithm shall be an interactive algorithm equivalent to ISOCIS, and shall process up to 16 channels of data, or any subset of channels defined by the user. The user shall be able to define such subsets on a per segment, or per aggregation of segments basis. All segments to be classified as an aggregation will have the same channel subset.

- A. The iterative and one-pass clustering algorithms shall each provide a capability to generate up to 60 clusters whose statistics can be saved for subsequent use in classification.
- B. The iterative clustering algorithm shall have the option of acquiring starting vectors from the one-pass algorithm in batch as well as interactive modes while retaining the current options. Provision for batch control of the parameters of the one-pass algorithm shall be made.
- C. The option shall exist to use either the L-1 or the L-2 distance for assignment of pixels to clusters. Specifically, these distances are:

$$1. \text{ (L-1) } D1(X,Y) = \sum_{p=1}^n |x_p - y_p|, \text{ and}$$

$$2. \text{ (L-2) } D2(X,Y) = \sum_{p=1}^n (x_p - y_p)^2$$

where n is the number of channels, x is the vector $\{x_1, x_2, \dots, x_n\}$, and y is the vector

$\{y_1, y_2, \dots, y_n\}$

- D. On a split iteration, the iterative clustering algorithm shall split clusters in order of their maximum standard deviation from largest to smallest. On a combine iteration, the iterative clustering algorithm shall combine pairs of clusters in order of the distance between them from smallest to largest. Any cluster can be split or combined only once per iteration.
- E. The option shall exist in clustering to correct the dynamic range of each Landsat acquisition in a composed image for the effects of sun angle, with either the L-1 or the L-2 distance. This correction is done by altering the distances with a diagonal matrix A . The new distances are then

$$D1S(x,y) = D1(Ax,Ay) = \sum_{p=1}^n A_{pp} |x_p - y_p|$$

and

$$D2S(x,y) = D2(Ax,Ay) = \sum_{p=1}^n A_{pp}^2 (x_p - y_p)^2$$

- F. Upon completion of the last iteration of the iterative clustering algorithm, all clusters with less than $P(N) = n + 1$ (where $n = \#$ channels) pixels shall be eliminated and those pixels assigned to the remaining clusters.
- G. The number of sets of default values for clustering parameters shall be increased from two to eight or more (with analyst option to override for each parameter). The strip generation parameter and the select statistics box parameter shall be removed from and the R2 parameter shall be added to the list of default parameters.

- H. The option to eliminate DO (designated other) and/or DU (designated unidentifiable) fields from clustering is required. The option to input maximum and minimum radiance values in each channel and to exclude from clustering those pixels outside the interval (min, max) is also required.
- I. There shall be an option to label each of the final clusters with the category of the closest starting or labeling vector. The distance used shall be the same as that used in pixel assignment. There shall be an option in batch processing to input a label for a given cluster number or to delete a specific cluster's statistics from classification.
- J. There shall be an option to input a cluster distance threshold for use in Subclass Match. If any cluster is further from the nearest subclass than this threshold distance, a conditional cluster map will be generated in which each such cluster is assigned a color that indicates: (1) which subclass it was matched to, (2) that it exceeded the threshold distance, and (3) that a conditional cluster map will be generated every time a segment is clustered.
- K. The cluster chaining algorithm indicated in the flow charts shall be used to chain similar clusters together.
- L. The user shall have the capability to select arbitrary subsets of training segment channels for further processing, or to allow an automatic feature selection algorithm to select an optimum channel subset. Furthermore, he shall have the capability to associate any acquisition date of training segment data with any biological phase for classification, and he shall have the capability to utilize training data from any one of three segments(s) to classify any other segment.
- M. The user shall have the option to compute bias corrected category percentages from cluster populations. This correction shall be made automatically in the event that all clusters are given the same category label. Correction involves.

p = the number of categories to be bias corrected.

N_m = the number of pixels excluded from clustering by the min/max, radiance test.

N_{DO} = the number of pixels in DO fields

a_i = number of pixels assigned to cluster number i ,

s_q = {set of clusters that were labeled category C_q }

$N_q = \sum_{i \in s_q} a_i$ = Number of pixels assigned to all clusters labeled C_q ,

s = {set of all clusters with labels other than A}

$N = \sum_{i \in s} a_i$ = Number of pixels assigned to all clusters not labeled A.

$B = N + N_{DO}$ HS

N_{p+1} = the number of pixels assigned to the totality of clusters whose label is other than A or $C_1, C_2, \dots, C_p$.

$m_{q,r}$ = The number of bias correction dots labeled C_q but assigned to the totality of clusters labeled C_r for q and $r \leq p$.

ORIGINAL PAGE IS
OF POOR QUALITY

$m_{q,p+1}$ = The number of bias correction dots labeled C_q but assigned to the totality of clusters whose label is other than A or C_t for $t \leq p$.

$m_r = \sum_{q=1}^p m_{q,r} + m_{p+1,r}$ = The number of bias correction dots assigned to the totality of clusters labeled C_r .

m_{p+1} = The number of bias correction dots assigned to the totality of clusters labeled other than A or C_q for $q \leq p$.

$\alpha_{q,r} = \frac{m_{q,r}}{m_r}$ or $\alpha_{q,r} = 0$ if $m_r = 0$ for $q \leq p$ and $r \leq p+1$

$$P(C_q) = \left(\sum_{r=1}^{p+1} \alpha_{q,r} \frac{N_r}{B} \right) \times 100$$

For each $q \leq p$, V_q is a $p+1$ -dimensional vector whose r^{th} component is

$$V_{q,r} = \left[\frac{N_r}{B} \times 100 \right]^2 \frac{(\alpha_{q,r})(1-\alpha_{q,r})}{m_r - 1}$$

$$\text{VAR} \left[P(C_q) \right] = \left(\sum_{r=1}^{p+1} V_{q,r} \right) \text{ for } q \leq p$$

$$B_q = \frac{N_q}{B} \times 100 - P(C_q) \quad q \leq p$$

The following segment summary information shall be output from cluster bias correction to the dot summary report.

$P(C_1)$ through $P(C_p)$

N_1 through N_p , N_{p+1}

$\alpha_{q,r}$ for $q \leq p$, $r \leq p+1$

$V_{q,r}$ for $q \leq p$, $r \leq p+1$

B_q for $q \leq p$

$$\text{VAR} \left[P(C_q) \right] \text{ for } q \leq p$$

and

$$\left[B_q - Z \cdot \sqrt{\text{VAR} \left[P(C_q) \right]}, B_q + Z \cdot \sqrt{\text{VAR} \left[P(C_q) \right]} \right]$$

where $Z = 1.645$

4.1.1.5 Modification of statistics.— The capability to perform the following operations on class or subclass statistics shall exist:

- A. Sun angle corrections
- B. Mean level corrections

The user shall be able to select from the following options:

- A. No corrections. (This shall be the default option.)
- B. Sun angle correction only.
- C. Mean level adjustment only.
- D. Sun angle correction followed by mean level adjustment.

These options shall be available to the user on a per segment to be classified basis. Class and subclass statistics shall be available as a user optional output following the optional modification of statistics. In classifying a merged image, the user shall have the option to classify the training segments using unaltered statistics.

The output shall be available in the following forms:

- A. Line printer, microfiche, or microfilm. This output shall include means, covariance matrices, and correlation matrices.
- B. Magnetic tape. It shall be possible to input statistics at this point by use of magnetic tape. The above capabilities are operational.

4.1.1.5.1 Sun angle correction: These corrections shall be performed using coefficients obtained from a table organized by sun angle of acquisition of training segment versus sun angle of the equivalent acquisition of the recognition segment. Note that, since up to four different acquisitions of each segment are combined to create up to 16 channels of data prior to beginning pattern recognition, one to four pairs (of sun angle of training segment data acquisition versus sun angle of recognition segment data acquisition) exist. For each such pair a multiplicative and an additive correction exists for each Landsat band.

These corrections are represented by 4×4 diagonal matrices, A, and 4-element vectors, B. An A and a B exist for each sun angle of acquisition pair (i.e., SEA PT and SEA PR), and are denoted

A (SEA PT, SEA PR), B (SEA PT, SEA PR)

where

SEA PT = Solar elevation angle of acquisition P ($\bar{P}=1-4$) of training segment

SEA PR = Solar elevation angle of acquisition P ($\bar{P}=1-4$) of recognition segment

All off-diagonal elements of A are zero. The on-diagonal elements,

a_{ii} are decimal numbers X.XXX between 0 and 5.000.

The elements of B_i

b_i are decimal numbers XX.XX lying between -99.99 and 99.99.

i is a subscript indicating the Landsat spectral band for which the a_{ii} 's and b_i 's are valid.

The values of A and B shall be stored in a table extending from SEA = 20° to 70° for both training and recognition segment. The table shall be addressable at increments of 1°, although it is not necessary to store values of A and B in the table increments closer than 5°. Interpolation may be used within a 5° interval.

For a given data set, consisting of acquisitions M, N, O, P, form a sun angle correction matrix

$$AT = \begin{pmatrix} A(\text{SEA MT}, \text{SEA MR}) & (0) & (0) & (0) \\ (0) & A(\text{SEA NT}, \text{SEA NR}) & (0) & (0) \\ (0) & (0) & A(\text{SEA OT}, \text{SEA OR}) & \\ (0) & (0) & (0) & A(\text{SEA PT}, \text{SEA PR}) \end{pmatrix}$$

where each entry is a 4 × 4 matrix. Also form a sun angle correction vector

BT = (B(SEA MT, SEA MR), B(SEA NT, SEA NR),

B(SEA OT, SEA OR), B(SEA PT, SEA PR))

Then for each subclass, the sun angle corrected mean vector, μ_c , is formed from the original mean vector μ as follows:

$$\mu_c = \mu \times AT + BT$$

The sun angle corrected covariance matrix Γ_c is formed from the original covariance matrix Γ as follows:

$$\Gamma_c = AT \times \Gamma \times AT.$$

The above capability is operational.

4.1.1.5.2 Mean level adjustment: The system shall compute the mean vector in the training segments and in the segment to be classified over the MLA field(s) in each. If sun angle correction was requested, the system shall correct the MLA vector for the training segment as in the previous section. Then compute $\overline{\Delta\mu} = \overline{\mu}$ (training segment) - $\overline{\mu}$ (segment to be classified). Then for each class or subclass mean, $\overline{\mu}$ (class), compute the mean level corrected vector $\overline{\mu}'$ (class) as follows: $\overline{\mu}'$ (class) = $\overline{\mu}$ (class) - $\overline{\Delta\mu}$. The above functions are operational.

4.1.1.6 Feature selection requirements. - The system must provide an optional feature reduction procedure employing up to and including 16 channels (The default option is to by-pass feature reduction.) The separability measure criterion shall be the Bhattacharyya distance measure. The feature reduction algorithm must be able to process statistics for all the classes and subclasses defined within the data base for each segment.

The feature reduction procedure must be able to operate in the following modes:

Mode 1: This mode is the default option if the number of channels is over eight. Optimal channel subsets will be selected by a "without replacement" procedure until the subset separability is at least x percent of the total channel separability. As a user option, black-and-white (B&W) film may be produced for best single channel and/or color film may be produced for the best three channels.

Mode 2: This mode uses all the channels and is the default option if the number of channels is eight or less.

Mode 3: This mode uses a subset selected by the analyst.

Mode 4: This mode will compute Bhattacharyya distance separability for acquisitions and combinations of acquisitions.

A priori probabilities for use in the computation of Bhattacharyya distance shall be returned from the data base unless overridden by analyst input. Override may be accomplished in both batch and interactive modes.

Hard copy output products shall consist of the following:

- A. Number and identification of channels and the total number of channels available.
- B. Interclass, intraclass, other, and total separability as defined in the flow charts computed using all the channels and the extracted set of features.
- C. The ratios of all of the above computed using all the channels and the extracted or analyst selected set of features.

4.1.1.7 Classification.-- The classifier will accept or retrieve the following inputs for each segment to be classified. This capability is operational.

- A. Class or subclass statistics for each class (as modified by the feature extraction process).
- B. A priori probabilities for each category for the segment to be classified will be retrieved from the fields data base. The user will have the capability to override this value through category input or by selecting computation of proportional subclass a prioris.
- C. Threshold values for each category for the segment to be classified will be retrieved from the fields data base.
- D. DO and DU field definitions will be retrieved from the fields data base.
- E. Test and training field definitions will be retrieved from the fields data base.
- F. Channels to be used.
- G. Pixels from the segment to be classified.
- H. Thresholding should be implemented on a subclass basis with analyst-selected percent thresholding down to 0.1 percent in increments of 0.1 percent. Based on the number of pixels in the subclass training field and the number of channels used for classification, the computer will automatically determine the required numerical input assuming a Fisher's F distribution.

The default threshold value should be selected by the computer and be based on the single numerical entry per category which gives the most accurate threshold next to the single numerical entry per category which gives the most accurate threshold next to the single numerical entry per subclass model, i.e., a default value per category level which does not significantly change the level of thresholding.

The analyst should have the option to bypass the automatic thresholding and input values at the subclass level.

The threshold option could be used at any of the following entry points: (1) Class summary menu; (2) all classification image menus; (3) fields data base card interpreter; (4) process control card input.

- 4.1.1.7.1 Two types of classification will be performed as follows:
- A. Classification of training and test fields.
 - B. Classification of segments for acreage estimation.

In the classification of test and training fields, pixels falling in DO and DU fields will be classified as though they were ordinary pixels. In classification for acreage estimation, all points in DO fields will be classified as other and all points in DU fields will be classified as unidentifiable, unless they also lie in DO fields. Classification will assign pixel to a certain subclass and record that assignment in the Classification Work Data Set (Appendix B-13) along with an associated likelihood value. The density function to be used in the assignment decision involves the following contribution from a given subclass:

$$P(x, k, i, j) = \left(\frac{\sigma}{n_k m_i} \right) 2\pi^{-d/2} \left| \Gamma_{k,i,j} \right|^{-1/2} e^{-\frac{1}{2} (x - \mu_{k,i,j})^T \Gamma_{k,i,j}^{-1} (x - \mu_{k,i,j})}$$

where x is vector representation of the pixel to be classified

k is the category identifier

i is the class identifier

j is the subclass identifier

σ is the a priori probability corresponding to category k

n_k is the number of classes in category k

m_i is the number of subclasses in class i of category k

d is the dimension of pixel vector x

$\mu_{k,i,j}$ is the mean vector of subclass (k,i,j)

$\Gamma_{k,i,j}$ is the covariance matrix of subclass (k,i,j)

In the nominal default classification to the category level, the following sum would be computed for each category k :

$$P(x,k) = \sum_{i=1}^{n_k} \sum_{j=1}^{m_i} P(x,k,i,j)$$

The pixel would then be assigned to the category with the largest associated sum. Specifically, it shall be assigned to the subclass within that category which has the maximal $P(x,k,i,j)$. Stored with that subclass assignment shall be the following associated likelihood value to be used for thresholding:

$$\min \left\{ \left(x - \mu_{k,i,j} \right)^T \Gamma_{k,i,j}^{-1} \left(x - \mu_{k,i,j} \right), 255 \right\}$$

If this number is larger than the chi-square value, the pixel is classified into the category threshold.

For any category k for which classification to the class level has been specified, n_k sums of the form

$$P(x,k,i) = \sum_{j=1}^{m_i} P(x,k,i,j)$$

would enter the overall consideration for largest density sum as above. The specific subclass assignment and associated likelihood value would be derived in the same manner.

4.1.1.7.2 The classifier will have the following outputs for each recognition segment:

A. Classification Map

1. The classification map shall indicate wheat versus other versus thresholded versus unidentifiable, in a transparency form, at the same scale as LACIE PI imagery for the recognition segment alone.
2. The classification map shall indicate the boundaries and identities of all fields by a technique which does not prevent the user from ascertaining the classification of any pixels. It is permissible to use multiple sheets to accomplish this.

3. The grey shades or colors used in the classification maps shall be consistent by category between segments, yet retaining contrast. For example, wheat on Segment A should be represented the same as wheat on Segment B. CAMS shall have the capability to specify default values.
- B. 1. Threshold values used and actual number of pixels thresholded/class, category and subclass.
2. Date and user identification of input of threshold values used.
- C. 1. A priori probabilities used per category.
2. Date and user identification of input of a priori values used.
- D. Classes and subclasses used, source of each subclass, i.e., clustering, B-chaining, etc.
- E. (Transformed) band means and variances for each class or subclass.
- F. Classified percentages for each category $C_1, C_2, \dots, C_p$ to be bias corrected where:

C_q = name of category to be bias corrected, $1 \leq q \leq p$.

p = number of categories to be bias corrected.

N_{DU} = number of pixels in DU fields.

N_T = number of pixels thresholded in classification

N_A = number of pixels classified as Category A

$B = 22,932 - N_{DU} - N_T - N_A$ —

N_{DO} = number of pixels in all DO fields

A = the name of a category whose pixels are to be excluded from classification percentages. A shall be a user input with default A=X.

X - usually an unidentifiable category (clouds/cloud shadows).

N_q = the number of pixels classified in category C_q .

N_{p+1} = the number of pixels classified into a category other than the categories A, $C_1, C_2, \dots, C_p$.

$$\%C_q = \frac{N_q}{B} \times 100; \quad \%C_{p+1} = \frac{N_{p+1}}{B} \times 100; \quad \%DO = \frac{N_{DO}}{B} \times 100.$$

- G. 1.. $\frac{\text{Number of pixels in DO fields}}{196*117 - \text{number of pixels in DU fields} - \text{number of pixels in category unidentified}}$

2. Corrected percentages for each category $C_1, C_2, \dots, C_p$.

For an index, r , $1 \leq r \leq p+1$

$$P C_q = \left(\sum_{r=1}^{p+1} \alpha_{q,r} \frac{N_r}{B} \right) \times 100$$

where

$m_{q,r}$ = number of bias correction dots labeled C_q by the analyst and C_r by the classifier

$m_{q,p+1}$ = number of bias correction dots labeled C_q by the analyst and as some category other than $A, C_1, C_2, \dots, C_p$ by the classifier.

m_r = Number of bias correction dots labeled C_r by the classifier.

m_{p+1} = number of bias correction dots that were labeled by the classifier as some category other than $A, C_1, C_2, \dots, C_p$.

The summary shall be output such that a sequential listing of type 1 dots is output first, followed by all dots (other than type 1) classified C_1 .

Note, any bias correction dots that were thresholded in classification or that fall in DU areas shall be excluded from $m_{q,r}$.

$$\alpha_{q,r} = \begin{cases} m_{q,r}/m_r & \text{if one or more bias correction dots} \\ & \text{were classified } C_r \text{ for } q=1,2,\dots, p \text{ and} \\ & r=1,2,\dots, p \\ 0 & \text{if no bias correction dots were classi-} \\ & \text{fied } C_r \text{ i.e. } m_r=0 \\ m_{q,p+1}/m_{p+1} & \text{for } q=1,2,\dots, p \text{ and } r=p+1 \end{cases}$$

A dot processing summary shall be provided for each segment processed (whether classification is performed or not). In general, this summary shall contain the following information sequentially listed with respect to dot numbers:

Dot number

Dot pixel, row number

Dot classification category (as determined by classifier if classified)
 Dot cluster label (label of cluster to which dot was assigned by clustering algorithm)
 Dot category label (if specified by analyst)
 Dot type (S-starting vector, L-labeling vector, B-bias correction, DO, DU)
 Dot green number
 Dot brightness

The summary shall be output such that a sequential listing of starting and labeling dots is output first, followed by all dots (other than starting or labeling dots) classified C_1 then C_2 sequentially through C_p , followed by dots classified other than, A, C_1 , $C_2, \dots, C_p$, followed by DU, threshold, then DO.

In addition, the following segment summary information shall be output to the dot summary report.

$P(C_1)$ through $P(C_p)$

N_1 through N_p , N_{p+1}

$\%C_1$ through $\%C_p$, $\%C_{p+1}$, $\%DO$

$\alpha_{1,1}$ through $\alpha_{1,p+1}$, _____

$\alpha_{2,1}$ through $\alpha_{2,p+1}$

through

$\alpha_{p,1}$ through $\alpha_{p,p+1}$,

For each $q \leq p$, a $p+1$ dimensional vector V_q whose r^{th} component is

$$V_{q,r} = \left[\frac{N_r}{B} \times 100 \right]^2 \frac{\alpha_{q,r}(1-\alpha_{q,r})}{m_r - 1}$$

$$\text{VAR} \left[P(C_q) \right] = \left(\sum_{r=1}^{p+1} V_{q,r} \right),$$

$$B_q = \frac{N_q}{B} \times 100 - P(C_q)$$

and

$$B_q - Z \sqrt{\text{VAR } P(C_q)} , B_q + Z \sqrt{\text{VAR } [P(C_q)]}$$

where $z = 1.645$.

- H. Number of pixels in DU fields
196*117
- I. Source of training data:
 - 1. Training segments(s), date(s) of acquisition.
 - 2. Training fields.
 - 3. External statistics.
- J. Channels used.
- K. Channel map used to relate training segment channels to recognition segment channels.
- L. Identity of recognition segment.
- M. Acquisition dates of recognition and training segments, including all flaggings of marginal data.
- N. Whether run was semiautomatic or interactive.
- O. Playback of analyst semiautomatic mode inputs.
- P. Playback of analyst interactive mode inputs.
- Q. Identifying run number and date of run. This item shall appear on all individual products.
- R. Biological phase.
- S. A classification summary similar to the ERIPS classification summary, generalized to include subclasses, pixels classified into each subclass, pixels threshold for each subclass, for all test and training fields in recognition and training segment(s).
- T. As item S, except by class, and category.

The following reports shall be included with the classification report when applicable:

- U. A fields definition report similar to that generated by the ERIPS, but including subclass name for each field.
- V. A clustering report including channels used for clustering, cluster population, class matches, cluster names, a detailed clustering report, interdistance clustering report, intradistance clustering report, cluster parameters, definition of cluster maps generated for PFC and also output on microfiche.
- W. A chaining report including fields/clusters chained, an intradistance report (B-distance; initial and final), initial number of subclasses and subclass identities, final number of subclasses and subclass identities, number of pixels per chain.
- X. A feature selection report including channels available, channels selected, interclass separability, total separability with all channels, total separability with selected subset of channels.
- Y. User name.
- Z. For interactive runs, if interactive field changes are made, a complete definition of fields, a priori probabilities, and threshold values used for classification shall be reported. A field boundary overlay shall also be generated and referenced.

- AA. Diagnostics and incorrect inputs.
- BB. The capability should exist to print selected pattern recognition reports.

4.1.1.7.3 Classification output products shall be distributed as follows:

- A. One copy, located on microfiche, of each of items A.2 to BB above will be delivered to EOD for the data processing analyst associated with the recognition segment.
- B. Items A.1, B.1, C.1, F, G, H, I.1, I.3, K, L, M, N, Q, R, S, T, U, V, W, X, Y, Z, AA, and BB shall be stored in the ISRRS for each classification run; later runs shall not supersede earlier runs.
- C. The following parameters shall be stored on magnetic tape for each (batch) classification run.

For the recognition segment:

- 1. Sample segment number
- 2. DPR number
- 3. Run ID
- 4. Bhattacharyya separability
 - a. All channels
 - b. Channel subset used
- 5. External statistics tape number
- 6. Number of recognition segment subclasses
- 7. Number of merged image fields
- 8. Number of pixels in D0 fields
- 9. Number of pixels in DU fields
- 10. Date of classification
- 11. Number of clustering runs
- 12. Number of chaining runs
- 13. Number of acquisitions used for classification
- 14. Channels used in classification
- 15. Nominal biological windows
- 16. Segment type identifier (spring/winter)
- 17. Sample segment use indicator (training or not training)
- 18. Dates of acquisition on merged image
- 19. Segment number, segment type indicator, and acquisition dates for each training segment
- 20. Threshold and a priori values used in classification (by category)
- 21. Subclass classification results
 - a. Subclass name
 - b. Statistics manipulation flags
 - c. Pixels classified into subclass
 - d. Pixels threshold from subclass

For each field in segment:

- 22. Field name
- 23. Subclass name
- 24. Field type
- 25. Field population
- 26. Number of subclasses
- 27. Name of each subclass

ORIGINAL PAGE IS
OF POOR QUALITY

28. Mean level adjustment, Sun angle correction flags for that (sub)class
 29. Pixels classified into that subclass
 30. Pixels threshold from that subclass
 31. Means and standard deviations for each class
 32. Information from items 4.1.1.7.2 F, G, H, and Q
 33. Cluster match and distance information
- This tape shall be stored by ISRRS in the physical library for use by CAMS and CAS.
34. Acquisition separabilities
 35. Cluster option indicators
 36. Cluster/dot report
 37. Cluster bias correction information
 38. Cluster green number & brightness value for each acquisition used in clustering.

The analyst evaluation of the recognition segment shall be manually passed to CAS and stored in the physical library by ISRRS. The following evaluation codes shall be used:

Code 01: Not machine processed - clouds, haze, snow, etc.

Means that the segment cannot be processed through the system because clouds, haze, etc., make interpretation and analysis impossible.

Code 02: Not machine processed - confusion crops or other interpretation problems.

This code should be used when a segment cannot be processed because of interpretation difficulties, especially when confusion between wheat/small grains and other crops is such that a wheat/small grains estimate cannot be determined. The rules for such a decision are to be negotiated between CAMS and CAS and included in the CAMS Detailed Procedures.

Code 03: Bad data - due to technical problems, not reordered.

In this case the segment cannot be processed due to technical problems arising from unsatisfactory histogram, excessive scanline drop, etc. If the segment is not reordered, this code is used.

NOTE: If the segment is reordered, no code should be listed and no data sheet should be passed to CAS since this segment will again be sent through the CAMS system for evaluation. Any segment that cannot be processed due to clouds, etc., or technical problems, should be passed as 01 or 03.

Code 05: Not machine processed - dormancy.

In cases where recognition is a problem due to the crop being in a state of dormancy this code is used.

Code 07: Not machine processed - preemergence.

In cases where the acquisition is prior to the criteria established for fall wheat recognition (TBD) this code is used.

Code 09: Not machine processed - multiple acquisitions.

Code 09 was developed to take care of the problem of processing multiple acquisitions of a segment at the same time. When multiple acquisitions are available at one time, only one segment, which is determined to be the better acquisition for interpretation and analysis by the CAMS analyst, is selected for complete processing. Other acquisitions are listed as Code 09 to indicate that they had been reviewed but were not processed through the system. This code enables CAS to review the segment results and account for each acquisition and how it was evaluated by CAMS.

Code 10: Unsatisfactory - unsatisfactory results for segment.

This code is to be used for any acquisition that processed through the system and, based on CAMS evaluation procedures, designated unsatisfactory.

Code 12: Unsatisfactory - no significant change.

This code is used when the new acquisition is evaluated to have no significant change from the previous unsatisfactory evaluation for the segment.

Code 14: Unsatisfactory - rework, reevaluated segment.

This code is used when a segment which was previously passed to CAS is reworked.

Code 18: Unsatisfactory - machine processed multitemporal analysis.

This code is used when more than one acquisition date is used to produce an unsatisfactory proportion estimate.

NOTE: All acquisition dates used in processing should be listed on the CAMS Evaluation Form.

Code 20: Marginal - marginal results for segment.

This code is to be used for any acquisition that is processed through the system and, based on CAMS evaluation procedures, designated marginal.

Code 22: Marginal - no significant change.

This code is to be used when the new acquisition is evaluated to have no significant change from the previous marginal evaluation for the segment.

Code 24: Marginal - rework, reevaluated segment.

This code is used when a segment which was previously passed to CAS is reworked.

Code 28: Marginal - machine processed multitemporal analysis.

This code is used when more than one acquisition date is used to produce a marginal proportion estimate.

NOTE: All acquisition dates used in processing should be listed on the CAMS Evaluation Form.

Code 30: Satisfactory - satisfactory results for segment.

This code is to be used for any acquisition that is processed through the system and, based on CAMS evaluation procedures, designated satisfactory.

Code 32: Satisfactory - no significant change.

This code is to be used when the new acquisition is evaluated to have no significant change from the previous satisfactory evaluation for the segment.

Code 34: Satisfactory - rework, reevaluated segment.

This code is used when a segment which was previously passed to CAS is reworked.

Code 36: Satisfactory - less than 5%, manually (hand) counted.

This code is to be used for any segment in which the proportion estimate is manually counted rather than machine processed.

NOTE: This category would not be used if the segment was a rework segment. Code 34 should be used.

Segments in an agricultural area that have a 0% proportion estimate are to be designated Code 36.

Code 38: Satisfactory - machine processed multitemporal analysis.

This code is used when more than one acquisition date is used to produce a satisfactory proportion estimate.

Code 40: Segment is totally nonagricultural.

This code is used when the segment is evaluated as having no agriculture at all, i.e., no discernable field patterns.

NOTE: Segments in an agricultural area that have a 0% proportion estimate are to be designated Code 36.

The ISRRS shall be designed to store a total of 40,000 such reports; the storage available for any given segment shall be limited only by total available storage space.

Storage of TBD portions of TBD number of these reports for TBD period following end of crop year is required. (Category 3).

The above results shall be accessible by recognition segment number and by training segment number(s).

D. A paper printout of a report, in the format as described in Appendix A, will be delivered to EOD for the data processing analyst associated with the recognition segment.

4.1.1.8 Semiautomatic mode.- The referenced flow charts indicate the semiautomatic mode. The semiautomatic mode shall, as a minimum, support the following operations:

- A. Preprocessing.- An option to produce spectral plots at C&I loading (Ref. Sec. 4.1.5).
- B. Classification.- The inputs shall be the following:
 1. A collection of one or more recognition segments for classification processing - no default.
 2. A set of from one to three training segments, or external statistics for the entire collection of recognition segments defined in (1).
 3. Channel map(s) relating channels of each acquisition of designated training segment(s) to channels 1 through 16.
 4. Channel subset instructions for processing entire collection of recognition segments. This channel subset must exist for each recognition segment and training segment defined in (1) and (2) above.
 5. Statistics computation mode by class and by training segment.
 6. Instruction to perform sun angle correction for selected recognition segments. Default shall be no correction.
 7. Instruction to perform mean level adjustment for selected recognition segments. Default shall be no adjustment.
 8. Instructions to classify test and training fields in training segments with unaltered statistics.

- 9. Feature reduction mode(s). The default is to by-pass feature selection.
- 10. User name.
- C. Computation of statistics for external output. Inputs are identical output. Inputs are identical to (A2) through (A10).
- D. Updates of field definitions, a priori probabilities, threshold values.
- E. An option to produce spectral plots (ref. SEC. 4.1.5).

The above capabilities are optional.

4.1.1.9 Interactive mode.— The interactive mode is viewed as a facility for processing those segments for which satisfactory results cannot be achieved in the semiautomatic mode. It shall combine the features of the semiautomatic mode and the current ERIPS.

Override of previous results may be accomplished by analyst addition or redefinition of fields, changes in classification, preprocessing modifications to statistics, selection of channels, etc. This shall be accomplished by online interaction.

Interfaces between the interactive and semiautomatic modes shall be as follows:

- A. Access to a common data base. (Some means will be required to minimize impact of the semiautomatic mode to the interactive mode and vice versa when the two modes require access to two segments which may not be simultaneously accessible.)
- B. Statistics output from the interactive mode may be used as an input to the semiautomatic classifier.
- C. Output from the interactive mode may be reinput to update the ISRRS. The ISRRS shall indicate the mode of update in its records.

The interactive mode is operational.

4.1.1.10 Selection of training segments.— The system shall have a mode by which it assists the analyst in specifying the best training segment.

For manual training segment selection, access to the following is required:

- A. Acquisition status at JSC of all segments.
- B. Acquisition status at GSFC of all segments.
- C. Nominal acquisition dates for each segment.

4.1.2 Training

(Req'd by: CAMS; Cat. 1; Impl. Resp.: CAMS/LACIE AVSB)

It is required that a training program adequate to indoctrinate the available personnel with a broad understanding of remote sensing, wheat growing, and recognition processing, together with a detailed knowledge of the LACIE CAMS data system and the procedures for its use be planned and operated.

4.1.3 Procedures Development
(Req'd by CAMS: Cat. 1; Impl. Resp.: CAMS/LACIE ASVB)

Development of the following procedures is required.

- 4.1.3.1 Image interpretation procedure.- A procedure exists which allows a trained analyst interpreter, using imagery produced from Landsat MSS data, to:
- A. Determine and categorize those ground cover classes which can be distinguished from wheat and those which cannot.
 - B. Identify training fields for each of those classes by line and pixel coordinates.
 - C. Label those training fields in a manner suitable for entry into the ADP system.
 - D. Perform the necessary quality control associated with the above operations.
- 4.1.3.2 Signature extension partitioning procedure.- A procedure exists which allows a trained data analyst, using information characterizing local soil types, meteorology, farming practices, crop calendars and Landsat MSS data, to divide wheat growing regions into regions of nearly constant wheat signature and to associate regions of similar signatures, to designate a set of training segments, and to redesignate training segments as required.
- 4.1.3.3 Training segment selection procedure.- A procedure exists to govern the selection of a training segment to be utilized in the processing of a recognition segment, given the above signature extension partitioning, the acquisition histories of the involved segments, and recent meteorological data.
- 4.1.3.4 Automatic data processing procedure.- A procedure exists to govern the choice of options used in the classification of segments given the recognition segment, training segment, acquisition histories and ancillary data.
- 4.1.3.5 ADP evaluation procedure.- A procedure exists to be used in evaluating the results of ADP classifications to assess their accuracy.
- 4.1.3.6 Rework procedure.- A procedure exists to be used in reworking classification results of inadequate accuracy.
- 4.1.4 Green Number and Haze Correction Computations
The system shall have the capability to compute the green number of each dot for each acquisition and to compute a set of green index numbers and a haze parameter for each acquisition. The green number for a pixel is the value of the Tasselled Cap greenness for the dot minus the greenness of the soil line. These numbers will be used to construct analyst aids in dot labeling. The set of green index numbers are cumulative histogram values of the green number with anomalous pixels removed. They will be used to assess the condition of the vegetation in a segment on a segment-wide

basis. The haze correction parameter is a measure of the amount of haze that was present in the atmosphere over a segment when an image of that segment was acquired. These numbers will be used for R&D purposes. All these computations should be performed for each acquisition that is loaded into the LACIE Data Base with the results output to cards according to the format of Section 4.1.4 and the cards delivered to CAMS. Both the green number and haze computations use several standard operations that will be described first.

4.1.4.1 Recalibrate the Data and Rotate to the Tasselled Cap Coordinate System

$$\text{Let } \chi = \begin{bmatrix} \chi_1 \\ \chi_2 \\ \chi_3 \\ \chi_4 \end{bmatrix}_{4 \times 1} \quad \text{where } \chi_j = \text{relative radiance as provided by Landsat MSS band } j.$$

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}_{4 \times 1} \quad \text{where } y_j = \text{relative radiance as provided by Landsat-2 MSS band } j \text{ when processed at GSFC with the pre-launch values of calibration parameters}$$

$$R_k = \begin{bmatrix} r_{ij}^k \end{bmatrix}_{4 \times 4} \quad \text{and} \quad \begin{bmatrix} \beta_1^k \\ \beta_2^k \\ \beta_3^k \\ \beta_4^k \end{bmatrix}$$

R_k and b_k describe the transformation from case k to the Tasselled Cap coordinate system where:

$k = 1$ denotes the case of Landsat-2 MSS data that was processed at GSFC using the pre-launch values of the calibration parameters.

$k = 2$ denotes the case of Landsat-2 MSS data that was processed at GSFC using the post-mission values of the calibration parameters.

$k = 3$ denotes the case of Landsat-3 MSS data

$$y_j = \sum_{i=1}^4 r_{ij}^k \chi_i + \beta_j^k \quad j = 1, 2, 3, 4 \quad (y = R_k^T \chi + b_k)$$

4.1.4.2 Sun angle correction

Let θ_o = reference solar elevation

θ = solar elevation of the acquisition being processed

$$Z = \begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \\ Z_4 \end{bmatrix} \quad \text{where } Z_j = \left(\frac{\sin \theta_o}{\sin \theta} \right) y_j \quad j = 1, 2, 3, 4$$

The vector Z represents the Landsat data for a pixel that has been recalibrated to the Landsat-2 pre-launch calibration, sun angle corrected and rotated to the Tasselled Cap coordinate system. This is the data that will be used in the screening test and in the haze computation. The vector y represents the Landsat data for a pixel that has been recalibrated and rotated to the Tasselled Cap but not sun angle corrected. This is the data that will be used in the green number computation.

4.1.4.3 Screen out bad data, clouds, cloud shadows and water, etc.

The screening process consists of 6 tests; the first 5 are used in the haze correction computation and all are used in the green number computations.

Test 1 - A pixel is labeled as garbled data if

$$Z_4 > Z_{4,\max}$$

$$\text{or if } Z_4 < Z_{4,\min}$$

$$\text{or if } Z_3 - 0.09375 Z_1 > Z_{3,\max}$$

$$\text{or if } Z_3 + 0.18750 Z_1 < Z_{3,\min}$$

$$\text{or if } Z_2 + 0.1 Z_1 < Z_{2,\min}$$

$$\text{or if } Z_2 + Z_1/1.8 > Z_{5,\max}$$

$$\text{or if } Z_2 - Z_1/1.2 > Z_{6,\max}$$

Test 2 - A pixel is labeled as cloud if not labeled garbled and $Z_1 > Z_{c\max}$ and $Z_3 + 0.1 Z_1 < Z_{c\min}$

/

Test 3 - A pixel is labeled as diffuse cloud if not labeled garbled or cloud and

$$Z_1 > Z_{H,max}$$

and

$$Z_3 + Z_1/7 < Z_{H,min}$$

Test 4 - A pixel is labeled as water if not labeled garbled, cloud or diffuse cloud and

$$Z_1 < Z_{w1}$$

and

$$Z_2 + Z_1/16 < Z_{w21}$$

and

$$Z_4 < Z_{w4}$$

and

$$Z_2 + Z_4 < Z_{w24}$$

and

$$Z_2 + 0.5 Z_1 + Z_3 + 5 Z_4 < Z_{w2134}$$

Test 5 - A pixel is labeled as cloud shadow if not labeled garbled, cloud diffuse cloud or water and

$$Z_2 - 0.4 Z_1 - 0.6 Z_3 - 0.6 Z_4 > Z_{s2134}$$

and

$$Z_1 - 0.4 Z_2 < Z_{s12}$$

Test 6 - A pixel is labeled as too bright for use in the soil line computation if

$$Z_1 > Z_{2,max}$$

4.1.4.4 Given an acquisition, the green value of the soil line and the green index number for the acquisition are computed as follows:

Step 1: Select correct Tasselled Cap transformation and Screening Test parameters

Step 2: Recalibrate data and rotate to the Tasselled Cap coordinate system

Step 3: Apply Sun Angle correction

Step 4: Screen the data using Tests 1-6

Step 5: Construct the cumulative histogram for the green value without sun angle correction (y_2) for all pixels that pass the screening test, i.e.,

$G(I)$ is the number of "good" pixels with $y_2 \geq I$

Step 6: Find the green value of the soil line (IS) as the smallest y_2 such that at least m pixels have $y_2 \leq IS$ where $m = \frac{1}{100} \times \text{number of pixels that passed the Screening Test.}$

NOTE: IS can be negative; however, $y_2 \geq -32$ for all pixels that passed the Screening Test.

Step 7: S is the set {0,1,2,3,...,20,22,24,...,30,35,40,45,50}

Step 8: For $J \in S$, $GN(J) = G(IS+J)$

Step 9: $GIN(0) = \frac{GN(0)}{22,932} \times 100$

Step 10: $GIN(k) = \frac{GN(k)}{GN(0)} \times 100$ for $k \in S$

Step 11: Round each $GIN(k)$ to I2
If $GIN(k) \geq 100$, set $GIN(k) = 99$

4.1.4.5 Given a pixel with Landsat measurement vector χ in some acquisition for which the green value of the soil line is IS, the green number (n) of the pixel is computed as follows:

Step 1: Select correct Tasselled Cap transformation and Screening Test parameters

Step 2: Recalibrate Landsat data and rotate to Tasselled Cap coordinate system

Step 3: Compute green number for pixel

$$n = y_2 - IS$$

4.1.4.6 The haze correction parameter (γ) is computed as follows:

Step 1: Select correct Tasselled Cap transformation and Screening Test parameters

Step 2: Recalibrate Landsat data and rotate to Tasselled Cap coordinate system

Step 3: Apply sun angle correction

Step 4: Screen acquisition using Tests 1-5

Step 5: Compute segment haze diagnostic

$$\hat{\chi} = \frac{1}{m} \sum_S \chi_i$$

S = the set of m pixels that passed the Screening
Test $m \leq 22,932$

Step 6: Compute

$$a = \left[\sum_{i=1}^4 \alpha_i^2 (\hat{\chi}_i - \chi_i^*) r_{i3}^k \right] \frac{\sin \theta_o}{\sin \theta}$$

$$b = \left[\sum_{i=1}^4 \alpha_i (\hat{\chi}_i - \chi_i^*) r_{i3}^k \right] \frac{\sin \theta_o}{\sin \theta}$$

$$c = \left[\sum_{i=1}^4 \hat{\chi}_i r_{i3}^k \right] \frac{\sin \theta_o}{\sin \theta} - y^*$$

If $b^2 - 2ac > 0$, then $\gamma = 1000 \times \left(\frac{b}{a} - \frac{1}{a} \sqrt{b^2 - 2ac} \right) + 300$

If $b^2 - 2ac \leq 0$, then $\gamma = 1000 \times \frac{b}{a} + 300$

where

$$\chi^* = \begin{bmatrix} * \\ \chi_1 \\ * \\ \chi_2 \\ * \\ \chi_3 \\ * \\ \chi_4 \end{bmatrix} \quad \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} \quad \text{and } y^* \text{ are parameters.}$$

4.1.5 The system must contain the option to make scatter and trajectory plots. The default is to produce both types of plots. These plots must be output via line printer as follows:

- 4.1.5.1 Trajectory plots.- At C&I loading time produce trajectory plots of green number vs. time and brightness vs. time for all 209 dots and all acquisitions on the data base. These plots are to be produced sequentially by dot number with three dots per plot. The brightness plot for a given set of three dots shall be directly beneath the green number plot for the same dots and there shall be three sets of plots for each page of printout. Each page should include segment number and acquisitions on the first line. The plots are to show a symbol 1, 2, 3 or * with green number (or brightness) read vertically and acquisition date horizontally. The symbols 1, 2 and 3 refer to the sequential order of the three dots represented (1 = first dot, 2 = second, 3 = third) and * records more than one dot with the same green number (or brightness) on a given acquisition. The green number axis shall range from 0 to 50 with 2 green numbers per line (25 lines). The brightness axis shall range from 30 to 100 with 2 brightnesses per line (35 lines). The horizontal axis shall consist of 42 spaces, with each space representing a nine day interval starting on September 1. Consecutive days shall be plotted adjacent to one another with the earliest date plotted as it falls, and the second day on its right. Each set of plots shall be annotated with dot number of the first dot in the sequence, its location (line and pixel) and a sequence of characters indicating the labels of the corresponding dots (a blank shall be output if no label exists on the data base). To allow approximate coordinates to be read off of either plot the vertical axes shall be annotated in increments of 10 units, and the time axis shall show the approximate start of alternate months starting with September.
- 4.1.5.1.1 Cluster trajectory plots.- Provide the option to produce trajectory plots from cluster means of the acquisitions used during batch processing. These plots are to be green number vs. time and brightness vs. time. They are to be produced sequentially by cluster number with three clusters per plot and arranged in the same format as the C&I trajectory plots (Ref. 4.1.5.1).
- 4.1.5.2 Scatter plots.- Provide the capability to produce scatter plots of green number vs. brightness on a fixed scale for each acquisition used in batch processing. There shall be two types of scatter plots:
- A. One plot per acquisition shall display the category label of each labeled dot. Unlabeled dots shall be displayed on this plot using a separate symbol. Multiple hits shall be recorded as the total number of hits rather than as a category symbol.
 - B. One plot per acquisition shall indicate the category in which each dot was classified. In the event that classification is not performed, the category label of the cluster to which the pixel is assigned should be used in place of the

classifier label. A separate symbol shall be used for dots that were thresholded or lie in DO/DU fields. Multiple hits shall be recorded as the total number of hits rather than as a category symbol.

The capability to relate dot spectral values to dot grid numbers shall be provided as follows. A background xy-grid shall be printed on each plot to allow the approximate coordinates of any dot to be read off. In addition, the following shall be provided:

Dot number table - A sequential listing by dot grid (not random sequence) number giving on the same row: dot grid number, dot grid coordinates (scan line number and sample number), analyst label, classifier label and dot green number and soil brightness for each acquisition rounded to the same precision as used in making the scatter plots.

Green number table - For each acquisition a listing of green number, brightness, dot number, grid intersection, analyst label, and classifier label. This list shall be ordered by increasing green numbers. Dots with the same green number shall be further arranged in order of increasing soil brightness.

The above plots and tables shall be annotated with segment number and acquisition date.

- 4.1.5.3 The option shall exist to apply sun angle and/or haze corrections to any of the above mentioned plots. The default option shall be to produce plots without these corrections.

SECTION 5.0
SUBSYSTEM INPUT REQUIREMENTS

Inputs required by CAMS from the other LACIE functional elements are designated in the following subparagraphs.

5.1 FULL-UP SYSTEM

5.1.1 DAPTS-Landsat Data Requirements

- 5.1.1.1 Regular Landsat imagery.— (Category 1; Implementation Responsibility: DAPTS) A 1:1 000 000 scale Landsat (9"x9") image of each Landsat frame used in LACIE is required four times each growing season. This imagery shall be CIR composite of bands 4, 5, and 7. The imagery should contain less than 20 percent cloud cover. The imagery should be representative of each of the four defined biological phases and for each of the 4800 LACIE segments.

These images are to be stored by the ISRRS. Access to approximately 40 frames/day is required.

- 5.1.1.2 Landsat MSS segment data.— (Category 1, except as noted; Implementation Responsibility: DAPTS) The CAMS performs classification on 117 line by 196 pixel segments of Landsat data. These segments are received from the Data Acquisition Preprocessing and Transmission Subsystem (DAPTS) via the ISRRS.

Segments shall be of two types, ordinary segments and designated training segments. The acquisition and preprocessing of ordinary segments is required in each of four time periods unless such acquisition is prevented by cloud cover, technical problems, or operational problems. The acquisition and preprocessing of designated training segments is required for each of up to 16 opportunities per year unless:

- A. Acquisition is prevented by cloud cover, technical problems, or operational problems.
- B. Acquisition occurs within two days after a previous successful acquisition.

The total number of such segments, up to 4800, shall be determined by CAS. Of these segments, no more than 960 may be designated training segments at any one time; however, the system shall be designed such that any segment may be made a designated training segment or ordinary segment at any time, as long as this does not result in more than 960 active designated training segments.

Preprocessing of segments shall consist of data quality evaluation, calibration, and registration.

~~ENDING PAGE BLANK NOT FOR~~

Data quality evaluation shall test the data for excessive cloud cover, loss of pulse code modulation decommutation synchronization, or noise. Data shall be evaluated as good, marginal, or bad. Good data will be input to the CAMS. Marginal data will be input to the ISRRS for CAMS, and replaced if good data are reordered in time. Bad data will not be input to the ISRRS.

Calibration shall adjust all data such that identical spectral radiances input to the Landsat MSS result in identical outputs of the GSFC preprocessing system. Outputs are regarded as identical if the noise introduced by the MSS-GSFC system is negligible in comparison with the noise introduced by atmospheric variations and other short-term temporal variations in the radiation input to the scanner.

Registration shall align each successive acquisition of a segment such that the RMS distance between the ground positions corresponding to a given image pixel on two separate passes is less than 1 pixel. Absolute registration accuracy to the surface of the earth shall be such that no segment is more than 5 statute miles from its nominal position.

When a sample segment maintains the same nominal position in two succeeding crop years, its actual position shall be registered within TBD pixels. (Category 1; Implementation Responsibility: DAPTS).

Capability to change any JSC defined parameters associated with a sample segment, or to order reacquisition of data, shall be provided.

5.1.1.3 PFC gain and bias factors.-- (Category 1; Implementation Responsibility: DAPTS) Gain and bias factors for each segment shall be computed in the preprocessing and included as part of the header record, for the purpose of controlling the gain and bias settings on the production film connector. These factors will be computed using an algorithm supplied by CAMS.

5.1.1.4 Mosaics.-- (Category 1; Implementation Responsibility: DAPTS) Black and white uncontrolled 1:1,000,000 scale mosaics of each of the LACIE country wheat growing regions are required. These mosaics shall be available before the start of analysis in any individual country. Each mosaic should be compiled using Band 5 (the red band) from imagery acquired during the wheat growing seasons. Two sets of each mosaic are required, one for the ADP country analysis group and one for the image analysis group. These shall be on a stable base.

For each major wheat growing region a north-to-south set of uncontrolled strip mosaics is required. This mosaic shall also be at a scale of 1:1,000,000, but should be compiled from a CIR composite of Landsat bands 4, 5, and 7. Each mosaic shall be compiled from a single Landsat pass acquired once during each biological

phase using an average or midlatitude of wheat region crop calendar to select the actual calendar time to make the data acquisition. Each image used should contain less than 20 percent cloud cover.

5.1.2 DAPTS-Field Data Requirements

Items A, B, and C of 5.1.2.1 are needed by the end of each calendar year for early evaluation of LACIE sites. Item C relates to bio-window 1 and data for fall planted crops are needed to coincide with FOD processing of Landsat data. These data will also support RT&E and accuracy assessment requirements.

5.1.2.1 Ground truth - CAMS.- (Category 1: Implementation Responsibility: DAPTS/S&AD-FSO/USDA)

DAPTS will provide the following data for each LACIE Intensive Test Site.

- A. One copy of a 1:24,000 scale rectified B&W photo-mosaic of each ITS. The photo-mosaic is to be made from high altitude aerial photography acquired once per year, just before wheat harvest.
- B. One copy of a 1:24,000 scale transparent overlay registered to the photo-mosaic of (A) above and delineating the boundaries of each field in the Test Site. Each field will be numbered and a tabular listing provided containing the following data:
 1. Field number
 2. Field acreage
 3. Crop and variety being grown
 4. Is field irrigated? If so, how?
 5. Is field fertilized? If so, how?
 6. Planting date

These data are to be acquired once per crop year, after spring planting and before wheat harvest.

- C. One copy of a 1:24,000 scale field boundary map registered to the photo-mosaic of (A) above and delineating the boundaries of only those fields containing fall planted crops in each Intensive Test Site in Texas, Kansas, Indiana and Washington. Each field will be numbered and a tabular listing provided containing the information of item B above. These data are to be acquired once per crop year, after fall planting and before beginning of dormancy.
- D. One copy of Periodic Observation data acquired at 18 day intervals from approximately 50 fields in each Intensive Test Site. Fields selected for observation must contain both wheat and other crop grown in the area. Fields selected must be 40 acres or larger in size. Adjacent fields planted in the same crop and divided by fences or farm roads may be grouped to obtain a 40 acre or larger area. Data are required from the end of dormancy for fall planted crops or spring crop plantings until harvest. Data are to be acquired concurrent with Landsat-2 overpasses. The following data are required in tabular form.

1. Growth stage
2. Percent ground cover
3. Plant height
4. Surface moisture
5. Weed growth
6. Field operations
7. Growth/yield detractants
8. Stand quality rating
9. Vertical color photos
10. Observation date
11. General comments on field and site conditions

5.1.3 DAPTS-Historical Agricultural Requirements

A. All of the following data requirements are needed to coincide with receipt of any Landsat data at JSC with the following exception: Where signature extension is being worked, the data are to be provided prior to Landsat acquisitions. The following is a list of countries which require data prior to Landsat acquisition:

Sept 1, 1975	U.S. Great Plains, Winter Wheat
Oct 1, 1975	USSR, Winter Wheat
Oct 1, 1975	Canada, Saskatchewan
Dec 1, 1975	Canada, Remainder
Jan 1, 1976	U.S. Great Plains, Spring Wheat
Feb 1, 1976	USSR, Spring Wheat.

5.1.3.1 Soil maps.— (Category 1; Implementation Responsibility: DAPTS/USDA)
Four copies of individual soil maps for the U.S. and each LACIE foreign country are required. The areas within each country for which soils information is required are those areas used for sample segment definition. Maps are to be provided as a positive print, at a scale of 1:5 000 000 or greater (as available), and at a known projection. The soil classification system preferred is the 7th approximation system, however the genetic (Dokuchev) system is also acceptable. In both classification systems there must be a sufficient level of detail on the maps provided to allow determination of the boundaries of soil suborders. Maps must be reviewed yearly and updated as new material becomes available prior to each LACIE phase.

5.1.3.2 Topographic maps.— (Category 1; Implementation Responsibility: DAPTS/FSO)
One copy of topographic maps for all areas containing sample segments are required for all LACIE countries. In the U.S. these maps are required in scales of 1:24 000 and 1:250 000 for each sample segment and 1:1 000 000 for areas containing one or more sample segments. Substitute maps of 1:62 500 scale are acceptable where 1:24 000 scale maps are not available. The contour interval shall be the standard for the scale map under consideration with the criterion being to show the greatest amount of detail while maintaining clarity.

For the LACIE foreign countries the maps required will be at, or as near as possible to, the scale and detail of the U.S. maps. In cases where topographic maps are not available, relief maps showing feature changes using hachures, tints or shading will be provided.

These maps (both U.S. and foreign) will be updated as better information becomes available.

- 5.1.3.3 A priori probabilities.— (Category 3; Implementation Responsibility: DAPTS/USDA) DAPTS has the responsibility of providing raw input data for the operation of a computer program to compute the prior probability of wheat in a given unit area. The input parameters are TBD.

A computer program specification, detailing the computational procedures to be employed, the frequency of computations, input data sources, output formats and any other required pertinent data and information required to develop the probabilities and their associated errors will be supplied to DAPTS by CAMS. This program will be implemented and operated by DSAD (Required by CAMS; Category 3; Implementation Responsibility: DAPTS/DSAD).

The various categories of agricultural ground cover to be considered are listed in the following subparagraph.

- 5.1.3.3.1 A priori probabilities: (Category 3; Implementation Responsibility: DAPTS/USDA) There is a requirement to determine that, if a segment of land is cropland, the cover class is a member of one or more of the following categories:

- A. Winter wheat
- B. Spring wheat
- C. Wheat
- D. Winter cereal
- E. Spring cereal
- F. Cereal
- G. Fall plowed crop
- H. Spring plowed crop
- I. Harvested winter wheat
- J. Harvested spring wheat
- K. Harvested spring cereal
- L. Harvested winter cereal
- M. Harvested fall plowed crop
- N. Harvested spring plowed crop

The estimated error in each of the above probabilities is required. These values are to be stored in the electronic ISRRS and are to be available to the CAMS whenever the CAMS is operating. An interactive query capability shall also be provided.

- 5.1.3.4 Ancillary summaries.— (Category 1; Implementation Responsibility: DAPTS/USDA) DAPTS shall output to the LPDL for the CAMS one copy of an ancillary data summary for each of the 4800 sample segments. The summaries shall contain information on soils, climate and

cropping practices as well as agricultural statistical data. The summaries shall be presented in the format shown by the attached examples (figs. 5-1 to 5-3). Each summary will be accompanied by its appropriate Locator-Map. Only two general formats are to be used - one for the U.S. and one for all foreign countries. The Soviet Union summary is an example of the format to be used in foreign countries. The text of these formats will be varied to meet the potential practices of a specific country.

The ancillary summaries are required at the county level for all counties in the United States containing one or more 5 x 6 n. mi. LACIE sample segments. For the foreign countries listed below, the data are required for political subdivision designated as acreage strata for aggregation purposes.

<u>Country</u>	<u>Acreage Strata</u>
Canada	Census Division
U.S.S.R.	Oblast
P.R.C.	Province
India	District
Australia	Division
Argentina	Partido
Brazil	Municipality

The agricultural statistical data required can be determined from the ancillary summary forms themselves. The cropping practices data requires an explanation of the particular information required.

Cropping practices are defined as any cultural activity that may create or alter the external morphological characteristics (appearance) of a field or crop signature response as detected by high altitude aircraft or spacecraft remote sensors. Examples of cropping practices are as follows:

- A. Cropping systems or crop rotation practices
- B. Tillage practices
- C. Irrigation practices
- D. Use of fertilizers
- E. Use of special-purpose crops
 - 1. Catch crops
 - 2. Cover crops
 - 3. Green-manure crops
 - 4. Soiling crops
 - 5. Silage crops
 - 6. Nurse crops
- F. Drainage practices
- G. Row directions and widths
- H. Chemical applications (herbicides, defoliants, etc.)
- I. Burning
- J. Field conditions post harvest (stacks, bales, plowed, etc.)
- K. Mowing patterns
- L. Grazing
- M. Harvesting practices

UNITED STATES ANCILLARY SUMMARY

County(ies) _____ State _____ Seg. No. _____
Seg. Coord. _____ CRD _____

Soil Type/Color _____ Climate (Köppen) _____
CROPPING PRACTICES Seasonal Precipitation
Annual Precipitation

Figure 5-1.- United States ancillary summary.

County/State

I	Land Use	Acreage (Most Recent Year)	Percent of Total County Area
	A. Total County Area		
	B. Total Cropland		
	1. Cropland Harvested		
	2. Cropland Pastured		
	3. All Other Cropland		
	C. Woodlands, Woodland Pasture		
	D. All Other Land		

E. Average Field Size (Acres) _____

F. Wheat (% Total of County Area) _____

II. Conservation Practices (Most Recent Year)

A. Irrigated Land (% County Area) _____

B. Contour (Acres)	Strip (Acres)	Terrace (Acres)
--------------------	---------------	-----------------

III. Crops and Agricultural Lands (Absolute Acreage and % of Total County Area)

Crops (Include All)	Most Recent Year			Second Most Recent Year		
	Acreage	Percent	Percent Irrigation	Acreage	Percent	Percent Irrigation
Wheat						

Figure 5-1.- Concluded.

SEGMENT NO. _____

SOVIET UNION
ANCILLARY SUMMARY

S.S.R. _____ OBLAST/KRAY _____

SEGMENT COORDINATES LAT. _____ LONG. _____

MAP NO. _____ SCALE _____

SOIL TYPE/COLOR _____ CLIMATE (KÖPPEN) _____

WMO NO. _____

I. CROPPING PRACTICES

Figure 5-2.- Soviet Union ancillary summary.

Segment No. _____

S.S.R. _____ Oblast/Kray _____

II.	Land Use	Area in 1000 Ha. Most Recent Year	Percent as Indicated
	A. SSR		
	B. Oblast/Kray		
	1. Cropped Area		
	a. Net area sown		
	b. Area sown more than once		
	c. Irrigated area		(% of 1A)
	2. Fallow Area		
	a. Fallow less than 1 year		
	b. Fallow 1-5 years		
	c. Fallow more than 5 years		
	3. Non-cultivated Area (a+b+c)		
	a. Forest		
	b. Permanent pasture/grazing		
	c. Other (marsh/desert/ urban, etc.)		

Comments: Figure 5-2.- Continued.

Segment No. _____

S.S.R. _____ Oblast/Kray _____

III. Agricultural Information

A. Average Field Sizes

B. Wheat (% of reporting Unit) _____

C. Conservation Practices

1. Irrigated land (% of reporting unit) _____

2. Contour (Acres) Strip (Acres) Terrace (Acres)

IV.

Crops Include A11	Most Recent Year 19			Remarks
	Area	Percentage	Percentage Irrigated	

Figure 5-2.- Concluded.

STATE _____
 CRD & No. _____

SEGMENT No. _____
 DATE _____

CROP CALENDARS
PERCENT OF AREA IN DEVELOPMENT STAGE BY SPECIFIED DATE

CROP	%AREA	MARCH	APRIL	MAY	JUNE	JULY	AUGUST	SEPTEMBER	OCTOBER
WINTER WHEAT	75 50 25								
BARLEY	75 50 25								
OATS	75 50 25								
RYE	75 50 25								
ALFALFA	75 50 25								
SORGHUM	75 50 25								
CORN	75 50 25								
SOYBEANS	75 50 25								
	75 50 25								
	75 50 25								
	75 50 25								
	75 50 25								

- YEARS OF DATA _____ - _____ AVERAGE
 (NO ADJUSTMENTS FOR WEATHER MADE)
- SOURCE OF DATA _____
- ESTIMATED ACCURACY OF DATES AS THEY WOULD APPLY TO ANY GIVEN YEAR
 ± _____ DAYS

Figure 5-3.- Crop calendars.

ORIGINAL PAGE IS
 OF POOR QUALITY

STATE, PROVINCE, OR OBLAST _____

SEGMENT NO _____

COUNTRY, _____

DATE _____

CROP CALENDARS
PERCENT OF AREA IN DEVELOPMENT STAGE BY SPECIFIED DATE

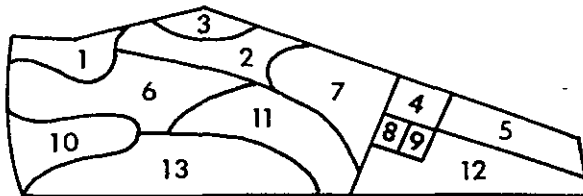
CROP	%AREA	JANUARY	FEBRUARY	MARCH	APRIL	MAY	JUNE	JULY	AUGUST	SEPTEMBER	OCTOBER	NOVEMBER	DECEMBER	JANUARY
	100													
	75													
	50													
	25													
	75													
	50													
	25													
	75													
	50													
	25													
	75													
	50													
	25													
	75													
	50													
	25													
	75													
	50													
	25													
	75													
	50													
	25													
	75													
	50													
	25													
	75													
	50													
	25													
	75													
	50													
	25													
	75													
	50													
	25													

SOURCE _____ ESTIMATED YEARLY SHIFT IN DATES $\pm$ _____ DAYS
 YEARS OF DATA _____
 ADJUSTED FOR WEATHER ____ YES ____ NO

Figure 5-3.- Concluded.

The following is a further amplification of the types of information required.

- A. Farm and field patterns: Data concerning the farm and field patterns in each wheat region will be required. In addition, representative schematics will be needed to illustrate "typical" field patterns and land uses in the wheat regions. The following schematic shows the land use on a farm. This illustrates the type of schematic required.



- B. Farm sizes: Information on the size of the farm and the individual fields (by crop type) is required, but a narrative explaining the data is highly desirable to the Image Interpreter. An example is this narrative excerpted from the Yearbook of Agriculture, 1958.

"The average commercial wheat farm in West-Central Kansas has about 715 acres of land. Of this, 525 acres are cropland and 190 acres are native pasture. About 330 acres of the cropland are seeded to wheat. As about two-thirds of it is seeded on summer-fallowed land, the average farm has about 140 acres of cultivated summer-fallow each year.

An average of about 40 acres, or one-sixth of the seeded wheat, fails to make a crop because of winter killing, drought or other hazard. If it rains after the wheat has failed some of this abandoned wheat land is planted to grain sorghum the following spring. The possibility of planting a catch crop of sorghum on abandoned wheat land is a unique characteristic of wheat-growing in the Southern Plains."

On the average, about 70 acres of grain sorghum and 40 acres of forage sorghum are planted each year. . .

The average farmer has about 40 head of cattle and no other livestock on a commercial scale."

- C. Crop rotation systems: Generalized information on the crop rotation systems will be needed. Three examples of this type of data are as follows:
1. Crop rotation systems in the Corn Belt:
 - a. 3-year rotation of corn, oats, red clover timothy sod
 - b. 4-year rotation of corn, soybeans, oats and clover
 - c. 5-year rotation of two years of corn, one year of oats, and two years of alfalfa-bromegrass.

2. One-year cropping systems:
 - a. Winter wheat followed by annual lespedeza
 - b. Winter rye grass alternated with a summer row crop such as corn.
3. Cropping practice in Yangtze Valley, China:
 - a. Cotton is hand broadcast in the wheat field one month prior to the wheat harvest.

Storage of these data on a sample segment basis in the LPDL is required. Approximately 20 such sets of data per day will be accessed.

- 5.1.3.5 Historical Crop Calendars.— (Category I; Implementation Responsibility: DAPTS/USDA) DAPTS is required to provide a mean historical crop calendar for each sample segment to assist the Analyst Interpreters (A.I.'s) in image interpretation. Each calendar shall describe the progress of wheat and all other crops which, taken together, constitute at least 95 percent of the cultivated area for which information is available. The year or years of historical data used in the development of the calendar shall be stated. Data sources shall also be stated. For foreign countries the calendars shall be prepared for areas similar in size to a U.S. Crop Reporting District (CRD). In the U.S. data will be required at the CRD level.

The calendars for each crop will be presented as percent area in a specified growth stage at a specified date. An example of the desired data product output format is attached.

- 5.1.3.6 Acquisition windows.— (Category I; Implementation Responsibility: DAPTS) The CAMS is responsible for determining the actual Landsat MSS data acquisition phases based upon estimated best acquisition phases plus Landsat MSS data already acquired.

The designated acquisition windows are defined by the CAS in terms of crop growth stage. These windows then input to the Yield Estimation Subsystem (YES). Estimates of dates delineating the growth stages on the basis of crop growth models, historical meteorological data, and/or other information will be used. These estimates are transmitted, via the ISRRS, to the DAPTS for action and to the CAMS for concurrence. If the CAMS fails to concur, it must provide acquisition windows to replace those provided by the YES.

To accomplish the above, the CAMS requires the dates delineating all crop acquisition windows for all segments as transmitted from YES to DAPTS. The ISRRS will maintain accessible records of all such transmissions.

5.1.4 YES Requirements

5.1.4.1 Crop calendar updates.— (Category 1; Implementation Responsibility: YES) A crop calendar is required for each acreage strata for which a sample segment exists. The following defines a crop calendar as required to support the image interpretation.

A "point plot curve" (from DAPTS) providing the expected time periods of the seedbed preparation, planting, phenological stages, and harvest of the significant crop types within a specified region for each biological year that LACIE is in operation. YES will provide a tabular update to these data every two weeks during the growing season.

For each field condition or biological stage the earliest date, the latest date, and that period in time when approximately 50 percent of a given crop has reached a defined stage should be included. (Ref. 5.1.3.5 for format.)

This type of calendar is required for wheat, its confusion crops (small grains), and other crops such that calendars are provided for 95 percent of the cultivated area in the neighborhood of the sample segment.

Updates to these calendars shall be provided to indicate early or late attainment of various growth stages of wheat and other crops in response to real-time meteorological data. The format shall be the same as used for historic calendars when more crops than wheat are updated.

Storage of these data on a sample segment basis in the ISRRS is required. Approximately 20 such sets of data per day will be accessed.

Individual crop calendars should be provided for each acreage strata for which a segment is designated for CAMS AI analysis. Sample segment numbers with country and locality descriptors should be included in the title data for each calendar.

Source data for the calendar construction should be clearly stated on each calendar, including the data of the latest meteorological adjustments applied.

Interpolated data should be distinguished from USDA or other statistical sources.

5.1.4.2 Meteorological summary.— (Required by CAMS; Category 1; Implementation Responsibility: YES) Meteorological data will also be utilized to provide weekly updates of the weather being experienced in specific regions. Such data will be similar to that being provided for the United States in the Weekly Weather and Crop Bulletin; i.e., normal moisture and temperature, wetter and warmer than normal, drier and warmer than normal, and drier and cooler than normal.

These data will be provided in tabulations which will apply to the sample segments. A narrative, which would flag areas where signature extension might not be successful due to some parameter other than temperature and/or precipitation, will also accompany the above. Such other parameters include wind damage, hail damage, or other anomalies occurring within the yield strata.

5.1.5 CAS Requirements

The following information is required at least one month prior to the start of any LACIE operational phase:

- 5.1.5.1 Sample allocation. - (Category 1; Implementation Responsibility: CAS) The nominal locations of all sample segments shall be indicated on transparent overlays suitable for registration upon USAF Operational Navigational Charts (ONC's) or Landsat uncontrolled mosaics. Sample segment center coordinates shall be latitude and longitude in degrees and minutes. These overlays shall also indicate the achieved Landsat ground tracks, LACIE frame boundaries, and LACIE overlap zones. Three copies of each such overlay are required.

- 5.1.5.2 Landsat acquisition rates. - (Category 1; Implementation Responsibility: CAS) Histograms indicating the expected acquisition rates in segments per month are required separately for winter and spring wheat for each country.

ALSO for each LACIE designated country provide individual segment numbers for each predicted acquisition date during the wheat phenological development stage.

- 5.1.5.3 Sample segment histogram. - (Category 1; Implementation Responsibility: CAS) Histograms indicating the distribution of number of sample segments per unit area for the spring and winter wheat regions of each country are required.

- 5.1.5.4 Flagged area output. - (Category 1; Implementation Responsibility: CAS) Sample segment classification that results in more than 20 percent change in the previous biophase percentage wheat area will be returned to CAMS for classification verification.

5.1.6 ISRRS Requirements

- 5.1.6.1 Initial imagery. - (Category 1; Implementation Responsibility: ISRRS) Initial imagery products shall consist of color composite and black and white imagery generated from Landsat segment data. This imagery shall be produced, at the nearest practicable scale, to 1:125 000, such that constant spacing between pixels occurs in the crosstrack direction and constant spacing between scan lines occurs in the along-track direction.

Black and white images shall display the value of each pixel in one channel on one acquisition as a set of levels of gray ranging from black to white. The number of shades of gray, n, shall be 8, 16,

or 256, as a user option. Data values shall be assigned to gray levels as follows:

1st gray level (black) corresponds to

$$y = 0 \text{ to } \left(\frac{256}{n} \right) - 1$$

2nd gray level corresponds to

$$y = \left(\frac{256}{n} \right) - 1 \text{ to } 2 \left(\frac{256}{n} \right) - 1$$

3rd gray level corresponds to

$$y = 2 \left(\frac{256}{n} \right) - 1 \text{ to } 3 \left(\frac{256}{n} \right) - 1$$

last gray level (white) corresponds to

$$y = (n - 1) \left(\frac{256}{n} \right) - 1 \text{ to } n \left(\frac{256}{n} \right) - 1$$

where

$$y = A (X + B) + C$$

where X is the pixel data value for the channel of interest, A is a scale factor, B is bias 1, and C is bias 2. A, B, and C may be used selected; default values are indicated below.

The user shall have the option to select any channel on any acquisition date for imagery generation; default values are indicated below.

Color images shall display the value of each pixel in a selected band in each of three colors (red, green, and blue). Each color shall be available in a range of brightness ranging from black to saturation of the film. The number of color levels, n, shall be the same for each color; the user shall have the option of selecting between n = 8, 16, or 256. Data values shall be assigned to each color level as follows:

1st color level (black) corresponds to

$$y = 0 \text{ to } y = \left(\frac{256}{n} \right) - 1$$

2nd color level corresponds to

$$y = \left(\frac{256}{n} \right) - 1 \text{ to } y = 2 \left(\frac{256}{n} \right) - 1$$

last color level (saturation of film) corresponds to

$$y = (n - 1) \left(\frac{256}{n} \right) - 1 \text{ to } n \left(\frac{256}{n} \right) - 1$$

where

$$y = A (X + B) + C$$

and X is the pixel data value in the selected channel, A is a scaling factor, B is the second bias factor, and C is the over-all bias.

A, B, and C may be user selected. Default values are indicated below.

The user shall have the option to assign any single channel/ acquisition date combination or pair thereof to any color in the image. Default values are indicated below.

All Landsat MSS imagery shall be generated at the same scale.

All default imagery products shall be produced within 24 hours of JSC receipt of the associated segment data. User requested imagery shall be produced within 24 hours of user request.

Flexibility to change default values under configuration control shall exist.

Gain and bias values for all Landsat MSS segments are supplied in the header record for all tapes received from GSFC on a per channel basis. These values shall be used as indicated below and shall be stored for future use.

All products shall have a superimposed grid at every 10th line and pixel starting at line one and pixel one. The grid shall not obscure more than one-half the area of any pixel.

The above capabilities are operational.

5.1.6.2 Landsat MSS tape products.- (Category 1; Implementation Responsibility: ISRRS) The products listed below are operational:

5.1.6.2.1 Screening film (Product Number 7): Screening film shall be produced for every acquisition of every segment. It shall be a moderate gain simulated color IR product using MSS bands 4, 5, and 7 and a monochrome moderate gain product for band 6.

MSS band 4 shall be assigned the color blue in the color composite; MSS band 5 shall be assigned the color green; MSS band 7 shall be assigned the color red; and MSS band 6 shall be used to create a monochrome image.

In each case, the overall bias, C, used for a band shall be taken from the tape header and multiplied by a TBD constant. Bias B shall be zero. The gain, A, shall be the value taken from the tape header multiplied by the above TBD constant.

- 5.1.6.2.2 Color composite positive (Product Number 1): This product shall be produced for the first acquisition in each biological phase of all currently designated training segments. It is a high gain simulated color IR product using Landsat MSS bands 4, 5, and 7.

The color blue shall be assigned to MSS band 4. The bias B shall be 0. The bias C shall be taken from the GSFC header record. The gain shall be taken from the GSFC header record.

The color green shall be assigned to MSS band 5. Gains and biases are as above.

The color red shall be assigned to MSS band 7. Gains and biases are as above.

- 5.1.6.2.3 Color composite positive/negative (Product Number 2): This product shall be produced for the same acquisitions as the color composite positive.

The color red shall be assigned to MSS band 5. The bias B shall be zero. Bias C shall be taken from the GSFC header. Gain shall be taken from the GSFC header.

The color blue shall be assigned to MSS band 6. The bias B shall be zero. Bias C shall be the bias value in the GSFC header. Gain shall be the gain value in the GSFC header.

The color green shall be assigned to MSS band 7. The gain and bias shall be as immediately above.

- 5.1.6.2.4 Black and white images (Product Number 4): A black and white image shall be produced for each band of the first acquisition in each biological phase of each currently designated training segment. These will be high gain images.

The biases B shall be zero. The biases C shall be obtained from the GSFC header. The gains shall be obtained from the GSFC header.

- 5.1.6.2.5 Temporal color composite: This product shall be produced upon request in each biological phase, if an earlier phase has been acquired. It is a high gain product allowing the interpreter to see temporal effects in a single image.

Eleven different acquisition sequences contain two or more acquisitions. Color and gain assignments for these sequences are TBD.

5.1.6.2.6 Image header: Each LACIE Landsat MSS image should be accompanied by the following header information on the film.

Job ID
System ID
Tape library
Date tape generated
Sensor
Orbit
Product requested
Sample segment ID, File number
Channels requested
Bias factors B
Landsat scene
PFC run date
PFC tape ID
Output mode
Number of pixel hits
Scaling factors A
Date of sensor data acquisition
PFC color balance factors
Color codes
Polarity
Repeat of pixel
Repeat of scan

5.1.6.3 ADP imagery.- (Category 1; Implementation Responsibility: ISRRS)
(See section 6.1.4)

5.1.6.4 Optional Landsat imagery.- (Category 3; Implementation Responsibility: ISRRS) TBD

5.1.7 Research Requirements
(Cat. 1; Impl. Resp.: Research) TBD

5.1.8 Test and Evaluation Requirements
(Cat. 1; Impl. Resp.: Test and Evaluation) TBD

SECTION 6.0
CAMS OUTPUT REQUIREMENTS

Outputs from CAMS which will be provided to the other LACIE functional elements or which are output by the CAMS and reinput to the CAMS are defined in the following subparagraphs.

6.1 FULL-UP SYSTEMS

6.1.1 Inputs to the ADP System
(Req'd by CAMS; Cat. 1; Impl. Resp. CAMS/NASA-DSAD)

- 6.1.1.1 Semiautomatic mode control inputs.- The Classification Applications Program semiautomatic mode requires control input. Inputs to the processing control data base should be by card. The analyst will process data in the semiautomatic mode by accomplishing the following functions via card input.
- A. Analyst name and run ID.
 - B. Recognition segment numbers.
 - C. Training segment numbers.
 - D. Channel maps between recognition segments and training segments.
 - E. Channels to be used for clustering (default is all).
 - F. Choice of one of two sets of clustering default parameters and option to select chaining for each cluster run.
 - G. Each cluster run initialized by input classes and subclasses with analyst defining the final class name for each clustering run.
 - H. Classes or subclasses to be used for each Bhattacharyya chaining run.
 - I. Normal computation of statistics by class option.
 - J. Utilization of external statistics from checkpoint tape.
 - K. Generation and writing of statistics on output tape.
 - L. Application of sun angle correction and mean level adjustment for various desired recognition-training segment pairs for subclass statistics. Including the option to classify the training segment with unaltered statistics.
 - M. Classes to be used in channel selection. Option to generate film. Option to distinguish between interclass distances and intraclass distances. Number of channels to be selected.
 - N. Option to set category thresholds and a priori probabilities for this job. Both classification and feature selection.
 - O. Definition of categories to be classified to class rather than category level.
 - P. Run purpose.
 - Q. A listing of the categories to be bias corrected and the name (if any) of the category to be excluded from bias correction.
 - R. Choice of whether or not to produce spectral plots and choice of type of trajectory plot to be produced. Defaults are to make spectral plots using all acquisitions in the Data Base; in the case of consecutive day acquisitions take the first acquisition. The default type of trajectory plot is option 1 (ref. SEC. 4.1.5).

S. Choice of whether or not to employ feature selection. The default shall be to by-pass feature selection unless a special input is made by the user.

- 6.1.1.2. Updates to fields data base.- The CAMS requires the capability to update the fields data base for each segment which contains the field definitions, a priori probabilities, and category threshold values for that segment. This is accomplished by means of the following interim products:
- A. Control cards containing field definitions, a priori probabilities, and/or threshold values (interim output product).
 - B. Tape containing field definitions, a priori values, and/or threshold values. This tape is generated by the interactive mode of the Classification Applications Program (interim output product).
 - C. Multifile tape containing field definitions. This tape is created by an offline program and unloaded from the data base. The segment information in the header of the tape is specified by the user or defaulted to the current segment identification.

- 6.1.1.3 Updates to dot data base.- The dot data base can be updated by means of the following products.
- A. Control cards containing dot number, type, and label.
 - B. Multifile tape containing dot number, type, and label. This tape is created by an offline program and unloaded from the data base. The segment information in the header of the tape is specified by the user or defaulted to the current segment identification.

- 6.1.2 Data Status Reports
(Req's by CAMS; Cat. 1; Impl. Resp.: CAMS/NASA-DSAD)

- 6.1.2.1 Indexed data report.- An indexed data report shall exist for every sample segment. This report shall be updated following any operation resulting in a change in, deletion of, or addition of parameters in the report or upon query.

This report shall indicate the following parameters for each sample segment:

1. Sample segment number
2. Country
3. Spring or winter wheat sample
4. Region, zone stratum
5. Latitude and longitude of nominal center point
6. GSFC registration and data quality parameters (reference scene, orbit, Landsat scene, percent cloud cover, registration-sharpness, P/B ratio, sun angle, sync-loss-lines, storage status, processing date)
7. List of acquisition dates for segment
8. Designated training segment or ordinary segment
9. Current definition of biological windows
10. History of updates to GSFC data base

11. Accession information for all data received from GSFC for this segment
12. Latest run ID.

- 6.1.2.2 Acquisition status matrices.- To accomplish manual training segment selection, access to the following items will be required:
1. Acquisition status at JSC of all segments.
 2. Acquisition status at GSFC of all segments.
 3. Nominal acquisition dates for each segment.

6.1.3 Classification Results

- 6.1.3.1 Classification summary report.- (Required by CAS, CAMS; Category 1; Implementation Responsibility: CAMS/NASA-DSAD) The processing of each segment by the classification processor in either the semi-automatic or interactive mode shall include a classification summary report. This report shall include the following:
- A. 1. Date and number of PFC classification map generated for this processing segment.
 - B. 1. Threshold values used and actual number of pixels thresholded/class, category and subclass.
 - C. 1. A priori probabilities used per category.
 - D. Classes and subclasses used on data control cards and for processing functions. Sources of all subclasses, i.e., clustering, B-chaining, etc.
 - E. (Transformed) band means and variances for each class or subclass.
 - F. Classified percentages for each category $C_1, C_2, \dots, C_p$ to be bias corrected where:

C_q = name of category to be bias corrected, $1 \leq q \leq p$.

p = number of categories to be bias corrected.

N_{DU} = number of pixels in DU fields.

N_T = number of pixels thresholded in classification

N_A = number of pixels classified as Category A

$B = 22,932 - N_{DU} - N_T - N_A$

N_{DO} = number of pixels in all DO fields

A = the name of a category whose pixels are to be excluded from classification percentages. A shall be a user input with default A=X.

X - usually an unidentifiable category (clouds/cloud shadows).

N_q = the number of pixels classified in category C_q .

N_{p+1} = the number of pixels classified into a category other than the categories A, $C_1, C_2, \dots, C_p$.

$$\%C_q = \frac{N_q}{B} \times 100; \quad \%C_{p+1} = \frac{N_{p+1}}{B} \times 100; \quad \%DO = \frac{N_{DO}}{B} \times 100.$$

- G. 1. Number of pixels in DO fields
 $196*117$ - number of pixels in category
 unidentifiable - number of pixels in
 DU fields
2. Corrected percentages for each category $C_1, C_2, \dots, C_p$.

For an index, $r, \quad 1 \leq r \leq p+1$

$$P C_q = \left(\sum_{r=1}^{p+1} \alpha_{q,r} \frac{N_r}{B} \right) \times 100$$

where

$m_{q,r}$ = number of bias correction dots labeled C_q by the
 analyst and C_r by the classifier

$m_{q,p+1}$ = number of bias correction dots labeled C_q by the
 analyst and as some category other than $A, C_1,$
 $C_2, \dots, C_p$ by the classifier.

m_r = Number of bias correction dots labeled C_r by the
 classifier.

m_{p+1} = number of bias correction dots that were labeled by
 the classifier as some category other than $A, C_1,$
 $C_2, \dots, p$.

The summary shall be output such that a sequential listing
 of type 1 dots is output first, followed by all dots (other
 than type 1) classified C_1 .

Note, any bias correction dots that were thresholded in
 classification or that fall in DU areas shall be excluded
 from $m_{q,r}$.

$$\alpha_{q,r} = \begin{cases} m_{q,r}/m_r & \text{if one or more bias correction dots} \\ & \text{were classified } C_r \text{ for } q=1,2,\dots, p \text{ and} \\ & r=1,2,\dots, p \\ 0 & \text{if no bias correction dots were classi-} \\ & \text{fied } C_r \text{ i.e. } m_r=0 \\ m_{q,p+1}/m_{p+1} & \text{for } q=1,2,\dots, p \text{ and } r=p+1 \end{cases}$$

A dot processing summary shall be provided for each segment processed (whether classification is performed or not). In general, this summary shall contain the following information sequentially listed with respect to dot numbers:

Dot number
 Dot pixel, row number
 Dot classification category (as determined by classifier if classified)
 Dot cluster label (label of cluster to which dot was assigned by clustering algorithm)
 Dot category label (if specified by analyst)
 Dot type (S-starting vector, L-labeling vector, B-bias correction, DO, DU)
 Dot green number
 Dot brightness

The summary shall be output such that a sequential listing of starting and labeling dots is output first, followed by all dots (other than starting or labeling dots) classified C_1 then C_2 sequentially through C_p , followed by dots classified other than, A, C_1 , C_2 , ..., C_p , followed by DU, threshold, then DO.

In addition, the following segment summary information shall be output to the dot summary report.

$P(C_1)$ through $P(C_p)$
 N_1 through N_p , N_{p+1}
 $\%C_1$ through $\%C_p$, $\%C_{p+1}$, $\%DO$
 $\alpha_{1,1}$ through $\alpha_{1,p+1}$,
 $\alpha_{2,1}$ through $\alpha_{2,p+1}$,
 through
 $\alpha_{p,1}$ through $\alpha_{p,p+1}$,

For each $q \leq p$, a $p+1$ dimensional vector V_q whose r^{th} component is

$$V_{q,r} = \left[\frac{N_r}{B} \times 100 \right]^2 \frac{\alpha_{q,r}(1-\alpha_{q,r})}{m_r - 1}$$

$$\text{VAR} \left[P(C_q) \right] = \left(\sum_{r=1}^{p+1} v_{q,r} \right),$$

$$B_q = \frac{N}{B} \times 100 - P(C_q)$$

and

$$B_q - Z \sqrt{\text{VAR} \left[P(C_q) \right]}, \quad B_q + Z \sqrt{\text{VAR} \left[P(C_q) \right]}$$

where $Z = 1.645$.

H. Number of pixels in DU fields
196*117

- I. Source of training data:
 - 1. Training segment(s), date(s) of acquisition.
 - 2. Training fields.
 - 3. External statistics.
- J. Channels used.
- K. Channel map used to relate training segment channels to recognition segment channels.
- L. Identity of recognition segment(s).
- M. Acquisition dates of recognition and training segments, including all flaggings of marginal data.
- N. Whether run in semiautomatic or interactive.
- O. Playback of analyst semiautomatic mode inputs.
- P. Playback of analyst interactive mode inputs.
- Q. Identifying run number, date of run. This item shall appear on all individual products.
- R. Biological phase.
- S. A classification summary similar to the ERIPS classification summary, generalized to include subclasses, for all test and training fields in recognition and training segment(s), and for the entire recognition training segment, giving pixels classified into each subclass.
- T. As item S, except by class.
- U. As item S, except by category.

The following reports shall be included with the classification report when applicable:

- V. A fields definition report similar to that generated by the ERIPS, but including subclass name for each field.
- W. A clustering report including channels used for clustering, cluster population, class matches, cluster names, a detailed clustering report, interdistance clustering report, intradistance clustering report, cluster parameters, definition of cluster maps generated for PFC and also output on microfiche.

- X. A chaining report including fields/clusters chained, an intra-distance report (B-distance; initial and final), initial number of subclasses and subclass identities, final number of subclasses and subclass identities, number of pixels per chain.
- Y. A feature selection report including channels available, channels selected, interclass separability, total separability with all channels, total separability with selected subset of channels, summary of film output of PFC.
- Z. User name.
- AA. For interactive runs, if interactive field changes are made, a complete definition of fields, a priori probabilities, and threshold values used for classification shall be reported. A field boundary overlay shall also be generated and referenced.
- BB. Diagnostics and incorrect inputs.
- CC. The user shall have the capabilities to print selected pattern recognition reports.

6.1.3.2 Electronic outputs.— Electronics outputs are as follows:

- 6.1.3.2.1 Ready access electronic outputs for sample segments: Items 1 to 30, defined in section 4.1.1.8.4, classification output products, as being saved on magnetic tape for each classification run shall be stored in the ISRRS physical library for ready access use by CAS.
- 6.1.3.2.2 Spectral value unload: The capability shall be provided to unload spectral values for all 209 grid intersection dots for any subset of the acquisitions on the data base to either tape or cards. Inputs shall be segment number, acquisition date(s), and output mode. Multiple segments may be output to one tape. If no acquisition date is input all acquisition dates are unloaded. Default output mode is tape. In tape output mode, the spectral information for the dots shall be output as an image in Universal format. In card output mode, the spectral information for the dots shall be output in a standard format along with the following parameters from the header record: segment number, acquisition date, solar elevation and the green value for the soil line in each acquisition.

6.1.4 ADP Imagery
(Req'd by CAMS; Cat. 1; Impl. Resp.: CAMS/ISRRS-DSAD)

Imagery products shall consist of transparencies at the approximate scale of 1:125 000. The exact scale shall be nominally identical to that of other LACIE applications imagery products. ADP image outputs shall be overlaid with a line and pixel ground identical to that used for initial imagery.

- 6.1.4.1 Field verification overlay.— (Product Number 12).— (Required by CAMS; Category 1; Implementation Responsibility: CAMS/ISRRS-DSAD)
This shall be an image indicating the boundaries and identities of all fields in a given segment. It shall be produced for a segment following any occasion on which the fields data base for that segment is updated, or following any interactive run in which field

definitions are changed prior to classification. This requirement is operational.

6.1.4.2 Cluster map (Product Number 9).-- (Required by CAMS; Category 1; Implementation Responsibility: CAMS/ISRRS-DSAD). Color coded conditional and unconditional cluster maps (ref. sec. 4.1.1.4) shall be output in image form each time a segment is batch processed. The boundaries and identities of all field(s) shall be indicated by a technique which does not obscure any pixels. It is permissible to utilize multiple sheets to accomplish this. This cluster map shall be produced for all clustering runs. This map should also contain an identifying number. This requirement is operational.

6.1.4.3 Classification map (Product Number 8).-- (Required by CAMS; Category 1; Implementation Responsibility: CAMS/ISRRS-DSAD) The classification map shall indicate wheat vs harvested wheat vs other vs unidentifiable vs thresholded. The boundaries and identities of all field(s) shall be indicated by a technique which does not obscure any pixels. It is permissible to utilize multiple sheets to accomplish this. The classification shall be produced for all classification jobs. This map should also contain an identifying run number. This requirement is operational.

The grey shades or colors used in the classification maps shall be consistent by category between segments, yet retaining contrast. For example, wheat on Segment A should be represented the same as wheat on Segment B. CAMS shall have the capability to specify default values.

6.1.4.4 Dot Overlay of Analyst Labels.-- The overlay shall indicate the current label of each analyst labeled dot and shall be at the same scale as the LACIE AI Imagery. It shall be produced for a segment following any occasion on which the dot data base is updated.

The segment summary information shall be output to the Dot Summary report whenever cluster bias correction is performed (paragraph 4.1.1.4.L).

6.1.5 Outputs to DAPTS
(Req'd by DAPTS; Cat. 1; Impl. Resp.: CAMS/LACIE ASVB)

6.1.5.1 Updates to GSFC data base orders (Computer Card Coding Sheets).-- CAMS shall implement the requests for changing the data acquisition and preprocessing order to GSFC. These changes will include specifying a need for additional processing of a sample segment for a given biological phase, or deleting sample segments, changing film processing flags, and changing color code information. The change will be specified by making the appropriate entries on a computer card coding sheet, using the format contained in the JSC card format documentation. These changes will then be submitted through the EOD Data Manager.

- 6.1.5.2 List of training segments.— CAMS shall specify, by segment number, those segments designated as training segments. This designation shall occur for up to 960 of the sample segments.
- 6.1.5.3 PFC gain and bias computation algorithm.— CAMS shall provide the algorithm to be implemented by GSFC in computing the gain and bias settings for the PFC.
- 6.1.6 Data Base Transaction Query Report
(Req'd by CAMS; Cat. 1; Impl. Resp.: CAMS/NASA-DSAD) A data base transaction report shall be generated for all alterations of stored field definitions, a priori probabilities, threshold values. This report may also be produced upon request to determine data base status. If the job produces a change in data base status, a copy will also be output for archival purposes. The report shall indicate the following:
- A. Run number
 - B. Date
 - C. Transactor name
 - D. Data base initial status (field names, classes, types coordinates; a priori probabilities; threshold values; date of most recent update).
 - E. Data base final status (as above) indicating all changes in status.
 - F. Accession information for field boundary overlay if applicable.
 - G. Accession information for all previous transaction reports.
- The total number of reports shall not exceed 100,000.
- 6.1.7 Resource Parameter Reports
- 6.1.7.1 Manual resources parameter reports.— (Required by SPE; Category 2: Implementation Responsibility: CAMS/NASA-ASVB) Throughput and resources information will be collected for SPE or an as required basis.
- 6.1.8 Spectral Plots
Each time a segment is batch processed spectral plots (ref. sec. 4.1.5) shall be output to a line printer if the spectral plot option has been chosen.

SECTION 7.0
INTERFACE REQUIREMENTS

Not applicable

RECEIVING PAGE BLANK NOT FILLED

SECTION 8.0

OPERATIONAL REQUIREMENTS AFFECTING THE SYSTEM DESIGN

The following requirements apply to all organizations implementing systems or procedures in support of CAMS.

- 8.1 **THROUGHPUT REQUIREMENTS** (Required by CAS; Cat. 1)
The system shall be capable of delivering to CAS the following number of satisfactorily processed segments after allowance is made for rework, accuracy assessment, quality control, data base maintenance, research, test, evaluation, and software development:
1 April 1975 - 31 December 1975 100 sample
 -- segments/week

1 January 1976 - 30 September 1976 170 sample
 -- segments/week

1 October 1976 - 31 March 1976 200 sample
 -- segments/week
- 8.2 **RESPONSE REQUIREMENT**
(Req'd by CAMS; Cat. 1)
The turnaround time for all products associated with a classification processing operation shall be not more than 24 hours.
- 8.3 **RELIABILITY REQUIREMENTS**
(Req'd by CAMS, Cat. 1)
The system reliability and throughput shall be such that the back of CAMS inputs or outputs accumulated due to failure of any system component is less than one day's peak processing capability.
- 8.4 **SECURITY REQUIREMENTS**
A LACIE Security Plan shall be prepared by each organization designated with implementation responsibility. The plan for each implementing organization shall define the specific measures that will be utilized by that organization to comply with the LACIE security requirements.

The LACIE security requirements shall be defined by the USDA and will be forwarded to the implementing organizations upon receipt by the LACIE Project Manager.

Each implementing organization shall submit their plan for approval to the LACIE Level III Change Board within 90 days after their receipt of the USDA requirements.
- 8.5 **DELIVERY OF PRODUCTS** (Req'd by CAMS; Cat. 1)
Delivery of all CAMS inputs and outputs to place of use is required. For those products to be analyzed by EOD personnel, the place of use is building 17, JSC.

- 8.6 QUALITY ASSURANCE REQUIREMENTS
(Req'd by CAMS; Cat. 1)
A LACIE Quality Assurance Plan is being prepared under the direction of the Quality Assurance Manager.
- 8.7 OTHER OPERATIONAL REQUIREMENTS
(Req'd by CAMS; Cat. 1)
The system shall be so designed that any segment can be processed on any day, regardless of the processing status of other segments.
- The system shall be designed to minimize the resources expended in preparing ADP inputs. Specifically, CAMS analysts shall not be responsible for card punching operations.

SECTION 9.0
CAMS VERIFICATION REQUIREMENTS

- 9.1 DATA SYSTEMS VERIFICATION
(Req'd by CAMS; Cat. 1; Impl. Resp.: CAMS/NASA-DSAD)
- 9.1.1 Design Reviews
Preliminary and critical design reviews of the CAMS data system shall be conducted. The CAMS manager shall have the authority to submit Review Item Dispositions against portions of the design affecting the CAMS.
- 9.1.2 Verification Testing
The verification of the CAMS data system shall be accomplished jointly by the EOD and the DSAD according to the following outline:
- A. The implementing organization shall prepare a test plan acceptable to the DSAD and to the CAMS manager.
 - B. Following acceptance of the test plan, the implementing organization shall prepare a test procedure acceptable to the DSAD and to the CAMS manager.
 - C. Following acceptance of the test procedure, the implementing organization shall perform the test, witnessed by the subsystem manager or his designate(s). A document of acceptable test completion shall then be prepared. Whenever practicable, acceptance tests will utilize data sets previously analyzed in the test and evaluation program.
 - D. Quality assurance will approve on test plans and test procedures
- 9.2 ANALYST CERTIFICATION
(Req'd by CAMS; Cat. 1; Impl. Response: CAMS/LACIE-ASVB)
- All analysts performing classification processing are required to be certified in accordance with the LACIE CAMS Analyst Certification Procedures Document, LACIE-00300.
- 9.3 PROCEDURES VERIFICATION
(Req'd by CAMS: Cat. 2; Impl. Resp.: CAMS/NASA-ASVB)
- 9.3.1 Operational Procedures
All operational CAMS procedures shall be documented and verified against their intended objectives, as stated in the Procedures Purpose. This verification shall be accomplished according to a plan and procedure acceptable to all organizations affected by the procedure to be verified.
- 9.3.2 Analysis Procedures
All CAMS analysis procedures shall be documented and tested before implementation. The documentation shall consist of a functional procedure which depicts the analysis in detail flow charts and a detailed procedure which technically describes the functional flow charts.

SECTION 10.0
RESEARCH REQUIREMENTS

(Required by CAMS; Cat. 4; Impl. Resp. Research)

TBD after development of LACIE Research, Test, and Evaluation Plan.

~~REPLACING PAGE BLANK NOT FILLING~~

SECTION 11.0
TEST AND EVALUATION

(Required by CAMS; Cat. 4; Impl. Resp.: Test and Evaluation)

TBD after development of LACIE Research, Test, and Evaluation Plan.

~~REPEATING~~ PAGE BLANK NOT PRINTED

APPENDIX A
REQUIREMENTS FOR
CAMS INTERFACE REPORT
AS PRODUCED FROM THE CAMS/CAS INTERFACE TAPE

A sample of the required report and an explanation of the data in each part of the report are attached.

Three options are required for the production of this report. They are as follows:

- A. The option to produce the entire report for all segments on a given tape,
- B. The option to produce the entire report for specific segments on a given tape, and
- C. The option to produce reports for all segments on a given tape, but for each segment, produce only the reports in Articles I, II, III and IV of the attached explanation and sample unless one of the computed PCC's for that segment is less than 80, and produce the entire report for that segment otherwise.

~~REPRODUCING PAGE BLANK NOT PERMITTED~~

EXPLANATION OF THE REPORT

I. HEADER INFORMATION. This includes

- A. DPAR number
- B. Record ID
- C. Segment number
- D. A list of the acquisitions used in processing
- E. The tape number from which the report was generated.

II. CLASSIFICATION AND CLASSIFICATION BIAS CORRECTION RESULTS. This includes

A. Type 1 Dot Label/Classification Matrix

This matrix shows the user assigned label and classified category of each starting and labeling dot. Results for each dot are displayed according to dot line and pixel number. The user label is to be printed first, followed by a slash and then the classified label. Dots that are thresholded in classification or that lie in DO or DU fields should have their classified categories reported as TH, DO, or DU accordingly.

B. Type 2 Dot Label/Classification Matrix

This matrix shows the user label and classified category of each bias correction dot and the classified category only for unlabeled dots. This information is to be reported in the same format as the Type 1 Label/Classification Matrix (Article II B).

C. Bias Correction Classification Report

This report lists:

- 1. For each category of interest
 - a. Category classified population
 - b. Category classified percentage
 - c. Category bias corrected percentage

d. Category variance

e. Category random sample estimate computed by

$$RSE() = \left(\frac{\text{the number of bias correction dots labeled in the specified category and classified in any category of interest (not thresholded)}}{\text{the total number of bias correction dots which were classified in a category of interest without thresholding}} \right)$$

$$\times 100 \times \left(\frac{22932 - \text{population DO-population unidentifiable}}{22932 - \text{population unidentifiable}} \right)$$

2. For DO, threshold and unidentifiable categories

a. total population of the category in question

b. the percentage of the category in question on the segment.

3. If W and S are both categories of interest, then compute and report items in Article II C.1 above for grain category as follows:

a. Population of grain category is the sum of the populations of category W and category S

b. Classified Percentage is the sum of the classified percentages of category W and category S

c. Corrected Percentage is

$$\alpha_G \left(\frac{N_W + N_S}{\text{Base}} + (1 - \alpha_N) \frac{N_3}{\text{Base}} \right) \times 100$$

where α_G = The number of bias correction dots labeled either W or S and classified (not thresholded) either W or S
The total number of bias correction dots labeled with any category of interest and classified (not thresholded) either W or S,

α_N = The number of bias correction dots labeled in any category of interest other than W or S and classified (not thresholded) in any category of interest other than W or S
 The total number of bias correction dots labeled with any category of interest and classified (not thresholded) in any category of interest other than W or S,

N_W is the population of category W,

N_S is the population of category S,

N_3 is the sum of the populations of all categories of interest other than W or S, and

Base is 22932 less the population of unidentifiable.

d. Variance is

$$\left[\frac{N_W + N_S}{\text{Base}} \times 100 \right]^2 \frac{\alpha_G(1-\alpha_G)}{n_W + n_S - 1} + \left[\frac{N_3}{\text{Base}} \times 100 \right]^2 \frac{\alpha_N(1-\alpha_N)}{n_3 - 1}$$

Where n_W is the number of bias correction dots classified (not thresholded) in category W,

n_S is the number of bias correction dots classified (not thresholded) in category S, and

n_3 is the number of bias correction dots classified (not thresholded) in a category of interest other than W or S.

e. Random Sample Estimate is

$$\frac{g_w + g_s}{n_w + n_s + n_3} \times \frac{N_w + N_s + N_3}{\text{Base}} \times 100$$

Where g_w is the number of bias correction dots labeled category W and classified in any category of interest and (not thresholded), and

g_s is the number of bias correction dots labeled category S and classified in any category of interest (not thresholded).

If any denominator is zero in these computations, the corresponding term should be zero.

D. Alpha Value Matrix.

The rows and columns of this matrix represent categories of interest. Each entry reports the Alpha value for the category of the row of the entry vs. the category of the column of the entry.

E. PCC's. These are computed as follows:

1. $PCC1 = \left(\frac{\text{Number of starting and labeling dots both labeled and classified (thresholded dots not included) in the same category of interest,}}{\text{Number of starting and labeling dots which were labeled and classified (thresholded dots not included) in any of the categories of interest.}} \right) \times 100$
2. $PCC2 = \left(\frac{\text{The number of bias correction dots both labeled and classified (excluding thresholded dots) in the same category of interest.}}{\text{The number of bias correction dots (excluding those thresholded) which were labeled and classified in any of the categories of interest.}} \right) \times 100$

If both categories W and S are categories of interest then compute PCC's for the grain category as:

3. $PCCG1 = \frac{\text{The number of starting and labeling dots labeled W or S and classified (not thresholded) either W or S + the number of starting and labeling dots labeled in a category of interest other than W or S and classified (not thresholded) in a category of interest other than W or S}}{\text{The number of starting and labeling dots which were labeled and classified (not thresholded) in a category of interest}} \times 100$

4. $PCCG2$ is computed the same way as $PCCG1$ except that bias correction dots are used rather than starting and labeling dots.

F. Bias Correction Vectors (Labeled: Classified).

This report lists for each ordered pair, (A, B), of categories of interest, the number of bias correction dots labeled in the first category (here category A) and classified in the second category (here category B). This number is reported as A:B. The report also includes the sum of all these entries as total.

G. The values D_w and D_s are computed as follows:

$$D_w = \sum_{p=1}^8 \sum_{q=1}^8 R_{pq} W_p W_q + \sum_{p=1}^8 T_p W_p \text{ or } 0 \text{ if category W not used}$$

$$D_s = \sum_{p=1}^8 \sum_{q=1}^8 U_{pq} S_p S_q + \sum_{p=1}^8 V_p S_p \text{ or } 0 \text{ if category S not used}$$

where $\{R_{pq}\}_{p=1, q=1}^8$, $\{T_p\}_{p=1}^8$, $\{U_{pq}\}_{p=1, q=1}^8$, and $\{V_p\}_{p=1}^8$ are constants (TBD).

W_1 = Bias corrected estimate for category W

S_1 = Bias corrected estimate for category S

W_2 = Machine estimate for category W

S_2 = Machine estimate for category S

W_3 = Random estimate for category W

S_3 = Random estimate for category S

W_4 = Variance of bias corrected estimate of W

S_4 = Variance of bias corrected estimate of S

$$W_5 = \frac{W:W}{W:S + S:S + N:S}$$

$$S_5 = \frac{S:S}{W:S + S:S + N:S}$$

$$W_6 = S_6 = \text{PCC1}$$

$$W_7 = S_7 = \frac{N:N}{W:N + S:N + N:N}$$

$$W_8 = S_8 = \text{PCC2}$$

The numbers W:W, W:S, etc. are as described in article II F above. If category W or category S is not a category of interest then each entry involving that category is assumed to be zero.

III. CLUSTERING AND CLUSTER BIAS CORRECTION REPORT

The format for this report is the same as the format for the bias correction report (article II). All entries are made using the category of cluster to which dots are assigned rather than classified label. Dots which are unassigned to a cluster are treated as thresholded dots.

IV. SEPARABILITY REPORT

This is a listing of feature selection Battacharyya separabilities for all available acquisition combinations.

V. CLASSIFICATION DATA

This reports the cluster name (subclass name), threshold value, and normalized apriori value for each cluster (subclass) used to train the classifier. It also lists the channels used in classification.

VI. CLUSTER REPORT

This report shows the following:

- A. Clusters Generated: The total number of clusters generated.
- B. Options Used: A listing of options used in clustering. Specifically sun angle normalization, DO/DU exclusion, or min/max radiance exclusion.
- C. For each of the clusters generated, a listing of cluster name, labeling dot, the clustering distance from the labeling dot to the cluster mean, the user assigned labels of dots assigned to that cluster, the cluster brightness and green number. This is followed by listing of the dots assigned to the cluster, their user assigned label and type and their clustering distance to the cluster mean. This list should be ordered by ascending clustering distance.

VII. STATISTICS REPORT

This report lists, for every subclass used to train the classifier, the subclass name, population, and mean and standard deviation in each channel.

VIII. DOT REPORT

This reports for each dot in the order present on the tape of the following information:

- A. Dot Number
- B. Line
- C. Pixel
- D. Dot type
- E. User Label
- F. Label of cluster to which dot was assigned
- G. Classification category (Th if thresholded)
- H. Brightness (For each acquisition)
- I. Green Number (used in classification)

I

SAMPLE

CAMS INTERFACE REPORT

ADPAR NO: # LA7-PC-PC01 A, CORDIN G15

E, TAPE NUMBER: 73240

S, RECORD ID # 164817 B 8001

D. ACQUISITION DATES

***** C SEGMENT NUMBER #1729

1	2	3	4
75304	00000	00000	00000

A-9

ORIGINAL PAGE IS
OF POOR QUALITY.

ORIGINAL PAGE IS
OF POOR QUALITY

A. TYPE1 DOT LABEL/CLASSIFICATION

PIXEL/ LINE	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190
10			W/C		N/S		W/W												
20	S/S			S/S	W/W													N/N	
30					S/S														S/C
40	W/W		N/N				S/S					N/DØ							S/S
50								N/N	N/N										
60	W/N															S/W			
70														N/N					N/N
80	X/W														S/W	W/S			S/V
90							N/DØ									W/W	W/W	S/S	N/N
100																	S/N	S/S	S/S
110						N/DØ		N/DØ				W/W	W/W		W/W	W/W	S/D	W/V	

B. TYPE2 DOT LABEL/CLASSIFICATION

PIXEL/ LINE	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190
10	/W	/W		W/S		W/S		/S	/W	/S	/S	/N	/N	/N	/W	/N	/N	/S	S/V
20		S/N	/W			W/W	/S	/S	/N	/N	/W	/S	/DØ	/DØ	/DØ	/DØ	N/DØ		/I
30	/W	/N	/N	S/N		/W	/N	/W	/W	/N	/TH	/DØ	/DØ	/W	/C	/DØ	/DØ	/DØ	
40		N/S		N/W	/S	/S		/S	/W	/N	/W		/DØ	/DØ	/N	/W	/C	/DØ	
50	/S	/S	/S	/S	/W	/N	/S			/W	/N	/DØ	/DØ	/W	/S	/N	/N	/S	/I
60		/N	/W	/S	/S	/N	/S	/S	/N	/S	/S	/W	/W	/N	/C		S/S	/W	/I
70	/C	/S	/W	/W	/C	/S	/S	/S	/DØ	/DØ	/S	/S	/W		/C	X/W	/C	/S	
80		X/W	/S	/N	/TH	/W	/W	/DØ	/DØ	/DØ	/S	/S	/S	/N			/N	/S	
90	/N	/N	/W	/N	/S	/S		/DØ	/S	/W	/W	/S	/W	/W	/W				
100	/W	/N	/S	/W	/S	/S	/DØ	/DØ	/W	/N	/S	/W	/W	/S	/S	/N			
110	/S	/N	/N	/N	/C	/S		/DØ		/N	/S	/N			/W				

C BIAS CORRECTION CLASSIFICATION REPORT

1.	CAT	CAT	CAT	CAT	CAT	CAT	CAT	CAT
	C	D	N	W	S	W	CAT	CAT
2. POPULATION	652	84	6099	206	6730	6202	0	0
3. CLASSIFIED %	2	0	26	0	29	27	0	0
4. CORRECTED %	0	0	16	0	43	23	0	0
5. VARIANCE %	0.00	0.00	37.55	0.00	37.55	55.82	0.00	0.00
6. RANDOM SAMPLE	0.00	0.00	15.84	0.00	31.67	23.75	0.00	0.00

2.	3.	4.	3.
CAT	CAT	CAT	CAT
DB	YH	UN	GRAIN
2754	205	205	12932
12	0	0	56
			71
			110.13
			2.55,43

	CAT	CAT	CAT	CAT	CAT	CAT	CAT	CAT
	C	D	N	W	S	W	CAT	CAT
7. CAT, 2F INT, /								
8. CAT, 2F INT,								
C	.000	.000	.000	.000	.000	.000		
D	.000	.000	.000	.000	.000	.000		
N	.000	.000	.000	.000	.000	.000		
W	.000	.000	.000	.000	.000	.000		
S	.000	.000	.000	.000	.000	.000		
W	.000	.000	.000	.000	.000	.000		

9. PCC1 = 62.5, PCC2 = 18.
10. PCCG1 = 72.5, PCCG2 = 45.

F BIAS CORRECTION VECTORS (LABELED CLASSIFIED)

CIC	0	CID	0	CIN	0	CIO	0	CIS	0	CIW	0
DIC	0	DiD	0	DIN	0	DiO	0	DIS	0	DiW	0
NIC	0	NiD	0	NIN	0	NiO	0	NIS	1	NiW	1
OIC	0	OiD	0	OIN	0	OiO	0	OIS	0	OiW	0
SIC	0	SiD	0	SIN	2	SiO	0	SIS	1	SiW	1
WIC	0	WiD	0	WIN	0	WiO	0	WIS	2	WiW	1
TOTAL #		9									

11. DW = (TBD)
12. DS = (TBD)

A-11

ORIGINAL PAGE IS
OF POOR QUALITY

TYPE1 DDT LABEL/CLUSTER ASSIGNMENT

PIXEL/ LINE	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190
10			W/C		N/S		K/W												
20	S/S			S/S	W/W													N/N	
30					S/S														S/S
40	W/W		N/N				S/S					N/N							S/S
50								N/N	N/N										
60	W/N															S/N			
70													N/N						N/N
80	X/W														S/C	W/W			S/W
90							N/N									W/W	W/W	S/S	N/N
100																	S/S	S/S	S/S
110							N/S		N/D				W/W	W/W		K/W	W/W	S/D	W/W

ORIGINAL PAGE IS
OF POOR QUALITY

TYPE2 DDT LABEL/CLUSTER ASSIGNMENT

PIXEL/ LINE	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190
10	/W	/W		W/S		W/S		/S	/W	/S	/S	/N	/N	/N	/N	/N	/N	/S	S/W
20		S/N	/W			W/W	/S	/S	/N	/N	/W	/S	/N	/S	/S	/N	N/N		/W
30	/W	/N	/N	S/N		/W	/N	/W	/W	/N	/W	/W	/N	/W	/W	/W	/W	/N	
40		N/S		N/W	/S	/S		/S	/W	/N	/W		/C	/N	/N	/W	/C	/N	
50	/S	/S	/S	/S	/W	/N	/S			/W	/N	/N	/W	/C	/S	/N	/N	/S	/S
60		/N	/W	/S	/S	/N	/S	/S	/N	/S	/S	/W	/W	/N	/C		S/S	/W	/W
70	/C	/S	/W	/W	/C	/S	/S	/S	/N	/W	/S	/S	/W		/C	X/W	/C	/S	
80		X/W	/S	/N	/N	/W	/W	/S	/N	/N	/S	/S	/S	/N			/N	/S	
90	/S	/S	/W	/N	/S	/S		/S	/S	/N	/W	/S	/W	/W	/W				
100	/W	/N	/S	/W	/S	/S	/S	/S	/W	/N	/S	/W	/W	/S	/S	/N			
110	/S	/N	/N	/N	/W	/N		/N		/N	/S	/N			/W				

A-12

BIAS CORRECTION CLUSTER REPORT

	CAT C	CAT D	CAT N	CAT S	CAT S	CAT W	CAT	CAT	CAT DZ	CAT UA	CAT UN	CAT GRAIN
POPULATION	936	114	7172	229	7981	6500	0	0	0	0	0	14481
CLASSIFIED %	4	0	31	0	34	28	0	0	0	0	0	62
CORRECTED %	0	0	28	0	38	26	0	0				69
VARIANCE %	0,00	0,00	73,65	0,00	73,65	90,21	0,00	0,00				286,53
RANDOM SAMPLE	0,00	0,00	13,04	0,00	17,39	13,04	0,00	0,00				35,43

ALPHA VALUE MATRIX										PCC1	PCC2
CAT, 1F INT, /	CAT C	CAT D	CAT N	CAT S	CAT S	CAT W	CAT	CAT		37,5	25,0
CAT, 2F INT,										PCCG1	PCCG2
										38,8	50,0
C	,000	,000	,000	,000	,000	,000					
D	,000	,000	,000	,000	,000	,000					
N	,000	,000	,000	,000	,000	,000					
S	,000	,000	,000	,000	,000	,000					
W	,000	,000	,000	,000	,000	,000					

BIAS CORRECTION VECTORS (LABELED CLASSIFIED)

CIC	0	CID	0	CIN	0	CIO	0	CIS	0	CIW	0
DIC	0	DID	0	DIN	0	DIO	0	DIS	0	DIW	0
NIC	0	NID	0	NIN	1	NIO	0	NIS	1	NIW	1
SIC	0	SID	0	SIN	0	SIO	0	SIS	0	SIW	0
WIC	0	WID	0	WIN	2	WIO	0	WIS	1	WIW	1

TOTAL = 10

SEPARABILITY REPORT

CHANNELS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	SEPARABILITY
AVAILABLE	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
1ACQ	X	X	X	X													4.12345
					X	X	X	X									1.23578
									X	X	X	X					3.45656
2ACQS	X	X	X	X	X	X	X	X									5.21765
	X	X	X	X					X	X	X	X					3.58029
					X	X	X	X	X	X	X						2.87599
3ACQS	X	X	X	X	X	X	X	X	X	X	X	X					0.52971

ORIGINAL PAGE IS
OF POOR QUALITYORIGINAL PAGE IS
OF POOR QUALITY

IV

CLASSIFICATION DATA

CLUSTER NAME THRESHOLD VALUE NORMALIZED
APRIORI VALUE

CoCL09	114,0	040
DoCL01	124,0	004
NoCL07	114,0	081
NoCL13	114,0	044
NoCL14	121,0	006
NoCL15	115,0	029
NoCL16	114,0	047
NoCL17	119,0	008
NoCL20	117,0	014
NoCL23	115,0	035
NoCL24	116,0	017
NoCL27	115,0	026
SoCL04	118,0	009
SoCL02	114,0	062
SoCL05	116,0	016
SoCL06	114,0	054
SoCL08	117,0	014
SoCL18	114,0	058
SoCL19	114,0	107
SoCL26	115,0	032
WoCL03	114,0	066
WoCL10	115,0	034
WoCL11	114,0	046
WoCL12	115,0	031
WoCL21	115,0	029
WoCL22	116,0	020
WoCL25	114,0	054

CHANNELS USED: 1 2 3 4

A-14

VI CLUSTER REPORT

A. CLUSTERS GENERATED: 27

B. OPTIONS USED:

		BRIGHTNESS		GREEN NUMBER	
CLUSTER NAME:	D0CL01	ACQUISITION 1	27	4	
LABELING D0T:	D0T206	ACQUISITION 2	0	0	
L1 DISTANCE:	2.30	ACQUISITION 3	0	0	
CATAGORIES:	D	ACQUISITION 4	0	0	

D0TS	DISTANCE	D0TS	DISTANCE	D0TS	DISTANCE	D0TS	DISTANCE	D0TS	DISTANCE
D 206 S	2.39								

		BRIGHTNESS		GREEN NUMBER	
CLUSTER NAME:	S0CL02	ACQUISITION 1	26	6	
LABELING D0T:	D0T023	ACQUISITION 2	0	0	
L1 DISTANCE:	2.47	ACQUISITION 3	0	0	
CATAGORIES:	S	ACQUISITION 4	0	0	

D0TS	DISTANCE	D0TS	DISTANCE	D0TS	DISTANCE	D0TS	DISTANCE	D0TS	DISTANCE
S 151	1.41	S 091	1.41	S 161	2.16	S 201	2.41	S 023 S	2.47
S 018	2.47	S 197 S	3.28	S 144	4.00	S 160	5.41		

		BRIGHTNESS		GREEN NUMBER	
CLUSTER NAME:	W0CL03	ACQUISITION 1	28	3	
LABELING D0T:	D0T206	ACQUISITION 2	0	0	
L1 DISTANCE:	1.61	ACQUISITION 3	0	0	
CATAGORIES:	W	ACQUISITION 4	0	0	

D0TS	DISTANCE	D0TS	DISTANCE	D0TS	DISTANCE	D0TS	DISTANCE	D0TS	DISTANCE
W 044	1.18	W 039	1.18	W 209 S	1.61	W 206 S	1.61	W 204 S	1.61
W 135 B	1.61	W 172	1.78	W 073	1.78	W 068	1.78	W 061 B	1.78
W 025 B	1.97	W 134 S	2.18	W 124	2.39	W 117	2.39	W 001	2.39
W 169 S	3.02	W 168 S	3.02	W 165	3.02				

A-15
ORIGINAL PAGE IS
OF POOR QUALITY

ORIGINAL PAGE IS
OF POOR QUALITY

CLUSTER REPORT

CLUSTER NAME:	00CL04	ACQUISITION	BRIGHTNESS	GREEN	NUMBER
LABELING DPT:	DPT199	1	26		2
L1 DISTANCE:	2.69	2	0		0
CATAGORIES:	0	3	0		0
		4	0		0

DPTS	DISTANCE	DPTS	DISTANCE	DPTS	DISTANCE	DPTS	DISTANCE	DPTS	DISTANCE
0 199 S	2.69								

CLUSTER NAME:	S0CL05	ACQUISITION	BRIGHTNESS	GREEN	NUMBER
LABELING DPT:	DPT188	1	18		2
L1 DISTANCE:	2.70	2	0		0
CATAGORIES:	S	3	0		0
		4	0		0

DPTS	DISTANCE	DPTS	DISTANCE	DPTS	DISTANCE	DPTS	DISTANCE	DPTS	DISTANCE
S 188 S	2.70	S 157	3.34	S 057 S	3.46	S 031	3.66		

CLUSTER NAME:	S0CL06	ACQUISITION	BRIGHTNESS	GREEN	NUMBER
LABELING DPT:	DPT043	1	26		4
L1 DISTANCE:	1.08	2	0		0
CATAGORIES:	S	3	0		0
		4	0		0

DPTS	DISTANCE	DPTS	DISTANCE	DPTS	DISTANCE	DPTS	DISTANCE	DPTS	DISTANCE
S 043 S	1.08	S 174	2.08	S 116	2.08	S 062	2.08	S 164	2.09
S 158	2.09	S 165	2.17	S 178	2.60	S 182	2.69	S 177	3.08
S 121	3.08	S 099	3.52	S 065	3.60				

CLUSTER NAME:	N0CL07	ACQUISITION	BRIGHTNESS	GREEN	NUMBER
LABELING DPT:	DPT060	1	33		4
L1 DISTANCE:	3.00	2	0		0
CATAGORIES:	N	3	0		0
		4	0		0

DPTS	DISTANCE	DPTS	DISTANCE	DPTS	DISTANCE	DPTS	DISTANCE	DPTS	DISTANCE
N 092	2.41	N 156	2.62	N 104	2.62	N 101	2.62	N 045	2.62
N 041	2.62	N 173	3.19	N 111 S	3.23	N 051	3.23	N 147	3.41
N 138	5.00								

A-16

CLUSTER REPORT

CLUSTER NAME:	S0CL08	ACQUISITION 1	BRIGHTNESS	29	GREEN	5
LABELING DOT:	D0T020	ACQUISITION 2	0	0		
L1 DISTANCE:	1.28	ACQUISITION 3	0	0		
CATEGORIES:	S	ACQUISITION 4	0	0		

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
S 011	0.99	S 076 S	1.28	S 020 S	1.28	S 004 B	2.28		

CLUSTER NAME:	C0CL09	ACQUISITION 1	BRIGHTNESS	30	GREEN	4
LABELING DOT:	D0T148	ACQUISITION 2	0	0		
L1 DISTANCE:	1.54	ACQUISITION 3	0	0		
CATEGORIES:	C	ACQUISITION 4	0	0		

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
C 148 S	1.54	C 090	1.54	C 131	1.86	C 119	1.86	C 115	1.86
C 003 S	1.86	C 129	2.59	C 074	2.59	C 110	3.46	C 070	3.46

CLUSTER NAME:	W0CL10	ACQUISITION 1	BRIGHTNESS	29	GREEN	5
LABELING DOT:	D0T024	ACQUISITION 2	0	0		
L1 DISTANCE:	2.61	ACQUISITION 3	0	0		
CATEGORIES:	W	ACQUISITION 4	0	0		

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
W 030	1.55	W 180	1.91	W 203 S	2.61	W 152 S	2.61	W 050	3.10

CLUSTER NAME:	W0CL11	ACQUISITION 1	BRIGHTNESS	30	GREEN	4
LABELING DOT:	D0T207	ACQUISITION 2	0	0		
L1 DISTANCE:	1.29	ACQUISITION 3	0	0		
CATEGORIES:	W	ACQUISITION 4	0	0		

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
W 207 S	1.29	W 081	1.29	W 195	1.85	W 139	1.85	W 053	1.85
W 002	2.48	W 140	2.55	W 054	3.29	W 118	4.48		

ORIGINAL PAGE IS
OF POOR QUALITY

CLUSTER REPORT

CLUSTER NAME	WDC12	ACQUISITION	BRIGHTNESS	GREEN NUMBER
LABELING DDT	DDT024	ACQUISITION 2	0	0
L1 DISTANCE	1.62	ACQUISITION 3	0	0
CATAGORIES	W	ACQUISITION 4	0	0

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
W 089	1.62	W 024 S	1.62	W 183	2.03	W 205	2.06	W 022	2.06
W 184	2.47	W 175	2.94	W 066	3.94				

CLUSTER NAME	NOCL13	ACQUISITION	BRIGHTNESS	GREEN NUMBER
LABELING DDT	DDT060	ACQUISITION 2	0	0
L1 DISTANCE	1.80	ACQUISITION 3	0	0
CATAGORIES	N	ACQUISITION 4	0	0

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
N 072	1.36	N 071	1.89	N 060 S	1.80	N 096 S	2.08	N 192	2.12
N 016	2.52								

CLUSTER NAME	NOCL14	ACQUISITION	BRIGHTNESS	GREEN NUMBER
LABELING DDT	DDT171	ACQUISITION 2	0	0
L1 DISTANCE	3.07	ACQUISITION 3	0	0
CATAGORIES	N	ACQUISITION 4	0	0

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
N 171 S	3.07								

CLUSTER NAME	NOCL15	ACQUISITION	BRIGHTNESS	GREEN NUMBER
LABELING DDT	DDT005	ACQUISITION 2	0	0
L1 DISTANCE	22.90	ACQUISITION 3	0	0
CATAGORIES	N	ACQUISITION 4	0	0

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
N 123	3.92	N 067	5.10	N 048	6.08	N 013	7.51	N 040	9.90
N 014	9.90	N 137	11.19						

CLUSTER REPORT

CLUSTER NAME:	NOCL16	ACQUISITION 1	BRIGHTNESS	GREEN	NUMBER
LABELING DDT:	D0T084	ACQUISITION 2	22	3	
L1 DISTANCE:	1.92	ACQUISITION 3	0	0	
CATAGORIES:	N	ACQUISITION 4	0	0	

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
N 093	0.89	N 088	1.40	N 150	1.89	N 142	1.92	N 084 S	1.92
N 017 S	2.08	N 087	2.25	N 200	2.41	N 028	2.41	N 021 B	3.69

CLUSTER NAME:	NOCL17	ACQUISITION 1	BRIGHTNESS	GREEN	NUMBER
LABELING DDT:	D0T069	ACQUISITION 2	14	0	
L1 DISTANCE:	1.54	ACQUISITION 3	0	0	
CATAGORIES:	N	ACQUISITION 4	0	0	

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
N 069 S	1.54	N 109	2.42	N 035	2.83	N 032	4.54		

CLUSTER NAME:	SOCL18	ACQUISITION 1	BRIGHTNESS	GREEN	NUMBER
LABELING DDT:	D0T170	ACQUISITION 2	24	5	
L1 DISTANCE:	1.99	ACQUISITION 3	0	0	
CATAGORIES:	S	ACQUISITION 4	0	0	

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
S 103	1.38	S 033	1.68	S 170 S	1.99	S 094	1.99	S 034	2.11
S 141	2.20	S 080	2.20	S 078	2.20	S 120	2.28	S 059 B	2.62
S 190 S	2.72	S 095	2.99	S 063	3.01	S 105	3.59	S 126	3.62

CLUSTER NAME:	SOCL19	ACQUISITION 1	BRIGHTNESS	GREEN	NUMBER
LABELING DDT:	D0T064	ACQUISITION 2	36	4	
L1 DISTANCE:	3.25	ACQUISITION 3	0	0	
CATAGORIES:	S	ACQUISITION 4	0	0	

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
S 027	1.78	S 176	1.86	S 125	1.86	S 077	1.86	S 190	2.25
S 132	2.78	S 006 B	2.78	S 010	2.86	S 102	3.14	S 154	3.25
S 122	3.25	S 064 S	3.25	S 191	3.44	S 008	3.65	S 026	4.14
S 146	4.25	S 106	4.44	S 166	4.86	S 005 S	4.86	S 136	5.25
S 083	5.25	S 112 B	5.75	S 153	6.75				

ORIGINAL PAGE IS
OF POOR QUALITY

CLUSTER REPORT

CLUSTER NAME:	NOCL20	ACQUISITION 1	BRIGHTNESS	GREEN NUMBER
LABELING DDT:	DDT128	ACQUISITION 2	11	2
L1 DISTANCE:	3.29	ACQUISITION 3	0	0
CATEGORIES:	N	ACQUISITION 4	0	0

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
N 128 S	3.29	N 056	4.60						

CLUSTER NAME:	NOCL21	ACQUISITION 1	BRIGHTNESS	GREEN NUMBER
LABELING DDT:	DDT149	ACQUISITION 2	30	3
L1 DISTANCE:	1.80	ACQUISITION 3	0	0
CATEGORIES:	W	ACQUISITION 4	0	0

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
W 155	1.53	W 149 S	1.80	W 086	2.30	W 130 B	2.41	W 098	2.91
W 163	3.09	W 049	3.20						

CLUSTER NAME:	NOCL22	ACQUISITION 1	BRIGHTNESS	GREEN NUMBER
LABELING DDT:	DDT007	ACQUISITION 2	29	9
L1 DISTANCE:	4.20	ACQUISITION 3	0	0
CATEGORIES:	W	ACQUISITION 4	0	0

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
W 108	2.80	W 009	3.20	W 166	3.77	W 038	4.20	W 019 B	4.20
W 007 S	4.20	W 055	4.74	W 127	6.80				

CLUSTER NAME:	NOCL23	ACQUISITION 1	BRIGHTNESS	GREEN NUMBER
LABELING DDT:	DDT159	ACQUISITION 2	32	3
L1 DISTANCE:	2.35	ACQUISITION 3	0	0
CATEGORIES:	N	ACQUISITION 4	0	0

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
N 187	1.92	N 193	2.14	N 097	2.14	N 036 B	2.31	N 159 S	2.35
N 042 B	2.71	N 194	2.91	N 202	3.08	N 196	3.91	N 012	4.08

CLUSTER REPORT

CLUSTER NAME:	NOCL24	ACQUISITION 1	BRIGHTNESS 16	GREEN NUMBER 3
LABELING DOT:	DOT085	ACQUISITION 2	0	0
L1 DISTANCE:	2.37	ACQUISITION 3	0	0
CATEGORIES:	N	ACQUISITION 4	0	0

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
N 198	1.58	N 075	2.04	N 082	2.17	N 085 S	2.37	N 029	3.17
N 143	3.37								

CLUSTER NAME:	WOCL25	ACQUISITION 1	BRIGHTNESS 32	GREEN NUMBER 6
LABELING DOT:	DOT058	ACQUISITION 2	0	0
L1 DISTANCE:	2.26	ACQUISITION 3	0	0
CATEGORIES:	W	ACQUISITION 4	0	0

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
W 052	1.91	W 047	1.91	W 167	2.24	W 107	2.26	W 058 S	2.26
W 114	2.91	W 046	2.91	W 113	2.93				

CLUSTER NAME:	SOCL26	ACQUISITION 1	BRIGHTNESS 20	GREEN NUMBER 4
LABELING DOT:	DOT189	ACQUISITION 2	0	0
L1 DISTANCE:	1.81	ACQUISITION 3	0	0
CATEGORIES:	S	ACQUISITION 4	0	0

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
S 179	1.13	S 189 S	1.81	S 145	3.55	S 079	4.05		

CLUSTER NAME:	NOCL27	ACQUISITION 1	BRIGHTNESS 29	GREEN NUMBER 5
LABELING DOT:	DOT133	ACQUISITION 2	0	0
L1 DISTANCE:	1.16	ACQUISITION 3	0	0
CATEGORIES:	N	ACQUISITION 4	0	0

DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE	DOTS	DISTANCE
N 017	0.70	N 133 S	1.16	N 181	1.47	N 015	1.70	N 162	2.70

CLUSTERING CHANNEL LIST:

1 2 3 4

A-21

ORIGINAL PAGE IS
OF POOR QUALITY

ORIGINAL PAGE IS
OF POOR QUALITY

VII

STATISTICS REPORT

CHANNEL	SURCL DOCL01 (00114)		SUBCL SOCL02 (01442)		SUBCL WOCL03 (01519)		SUBCL BOCL04 (00229)		SURCL SOCL05 (00389)	
NUMBER	MEAN	ST DEVA	MEAN	ST DEVA	MEAN	ST DEVN	MEAN	ST DEVN	MEAN	ST DEVN
1	12.04	0.4	13.24	0.9	13.82	0.7	13.92	1.1	11.94	0.9
2	17.13	0.5	14.63	0.7	17.90	0.4	17.91	0.6	12.24	0.8
3	15.77	0.8	16.20	0.8	15.70	0.5	13.38	0.6	8.56	0.8
4	7.25	1.1	7.66	0.8	7.40	0.7	6.69	0.8	4.08	0.8

CHANNEL	SUBCL SOCL06 (01251)		SUBCL NOCL07 (01871)		SUBCL SOCL08 (00338)		SUBCL BOCL09 (00936)		SURCL WOCL10 (00797)	
NUMBER	MEAN	ST DEVN	MEAN	ST DEVN	MEAN	ST DEVN	MEAN	ST DEVN	MEAN	ST DEVN
1	13.96	0.6	15.71	1.1	12.96	0.2	15.13	1.3	14.00	0.4
2	15.76	0.9	20.11	0.7	18.10	0.7	18.56	0.8	17.10	1.1
3	14.66	0.8	20.19	0.8	17.62	1.0	16.94	0.6	17.63	0.8
4	6.46	0.7	8.60	0.7	8.32	0.6	8.98	0.4	6.92	0.3

CHANNEL	SURCL WOCL11 (01065)		SUBCL WOCL12 (00721)		SUBCL NOCL13 (01019)		SUBCL NOCL14 (00154)		SURCL NOCL15 (00680)	
NUMBER	MEAN	ST DEVN	MEAN	ST DEVA	MEAN	ST DEVN	MEAN	ST DEVN	MEAN	ST DEVN
1	15.40	0.8	16.30	1.0	16.78	0.9	8.23	1.0	22.76	3.2
2	18.15	0.7	18.28	1.0	18.28	0.7	6.87	1.5	28.41	3.7
3	17.72	0.8	16.96	0.9	20.20	0.8	3.19	1.7	29.09	3.7
4	7.98	0.3	7.00	0.5	8.36	0.9	1.98	1.0	11.70	1.3

CHANNEL	SURCL NOCL16 (01079)		SUBCL WOCL17 (00203)		SUBCL SOCL18 (01345)		SUBCL SOCL19 (02473)		SURCL NOCL20 (00327)	
NUMBER	MEAN	ST DEVN	MEAN	ST DEVA	MEAN	ST DEVN	MEAN	ST DEVN	MEAN	ST DEVN
1	12.76	0.9	11.81	1.1	12.60	1.0	18.11	1.1	11.78	0.9
2	13.76	0.7	10.64	1.0	13.86	0.8	21.45	1.4	9.86	0.8
3	12.07	0.9	4.38	1.1	14.15	0.8	22.24	1.8	1.29	1.0
4	5.66	0.7	1.63	0.8	6.69	0.8	9.46	0.8	0.35	0.5

CHANNEL	SURCL WOCL21 (00680)		SUBCL WOCL22 (00470)		SUBCL NOCL23 (00825)		SUBCL NOCL24 (00405)		SURCL WOCL25 (0124A)	
NUMBER	MEAN	ST DEVN	MEAN	ST DEVA	MEAN	ST DEVN	MEAN	ST DEVN	MEAN	ST DEVN
1	13.69	0.4	13.79	1.0	16.58	1.1	10.17	0.8	14.07	0.4
2	19.56	0.6	14.49	1.1	20.22	0.7	10.30	0.9	18.51	0.9
3	16.75	1.2	20.45	2.2	17.89	0.9	7.60	1.1	20.66	1.3
4	7.89	0.5	9.52	1.4	8.40	0.6	3.89	0.8	8.99	0.9

CHANNEL NUMBER	SUBCL S0CL26 (00743)		SUBCL N0CL27 (00609)	
	MEAN	ST DEVN	MEAN	ST DEVN
1	11.09	0.8	14.00	0.0
2	12.34	0.9	17.65	0.9
3	10.96	1.0	17.73	0.7
4	5.34	0.8	8.09	0.3

DOT REPORT

A DOT #	B LINE	C PIXEL	D TYPE	E LABEL	F CLUSTER	G CLASS	H BRIGHTNESS/	I GREEN #
1	10	10			N	N	61	8
2	10	20	L	N	N	N	54	7
17	10	170	B	S	W	S	36	0
...								
...								

A-23

ORIGINAL PAGE IS
OF POOR QUALITY

ORIGINAL PAGE IS
OF POOR QUALITY

Project Control Office
LACIE Document
Distribution List
September 22, 1978

September 22, 1978

SF4/R. G. Musgrove
R. L. Nance
B. E. Spiers
D. R. Thompson
T. T. White
LEC/L. F. Abbotts
J. G. Baron
M. L. Bertrand
B. L. Carroll (4)
W. J. Daley
T. B. Dennis
J. J. Dishler
C. H. Jeffress
C. H. Liszcz
D. R. Marquis
T. C. Minter
J. Morgan
C. T. Nosworthy
D. G. Saile
F. W. Solomon (4)
P. C. Swanzy
F. Tragni
T. B. Wilkinson
Aeroneutronic Ford
3-3A/G. Austin

USDA, Washington
S. Spornack
F. Warren

IBM-16
K. Blessitt

